23 May 2008 | S10

Science

Microbial Worlds

AAAS

COVER

A hot spring in Bali, Indonesia. The discovery of thriving microbial communities in such unexpected places has motivated investigation into the diversity and distribution of microbial life. The special issue beginning on page 1027 explores the microbial world.

Image: Sylvain Grandadam/Getty Images

DEPARTMENTS

EDITORIAL

SPECIAL SECTION

Microbial Ecology

INTRODUCTION

NEWS

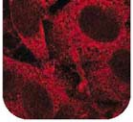
REVIEWS

>> Editorial p. 985; News stories pp. 1001 and 1006; Science Express Report by R. E. Ley et al.; Brevia p. 1046; Research Article p. 1047; Report p. 1081; for online content see p. 979 or go to www.sciencemag.org/microbialecology/

NEWS OF THE WEEK

NEWS FOCUS

CONTENTS continued >>



SCIENCE EXPRESS

www.sciencexpress.org

CELL SIGNALING

The Rag GTPases Bind Raptor and Mediate Amino Acid Signaling to mTORC1

Y. Sancak et al.

Nutrients, specifically amino acids, are sensed by small guanosine triphosphatases, which bind to a signaling complex, moving it close to the nucleus where it initiates cell growth.

10.1126/science.1157535

MICROBIOLOGY

Evolution of Mammals and Their Gut Microbes

R. E. Ley et al.

Genomic sampling of the microbes in the feces of 60 mammals shows that herbivores harbor the most diversity and that individuals of the same species have the same flora.

>> Microbial Ecology section p. 1027

10.1126/science.1155725

CELL BIOLOGY

β-Arrestin–Mediated Localization of Smoothened to the Primary Cilium

J. J. Kovacs et al.

β-arrestin, which has several known roles in signaling systems, also links a key receptor to a motor protein so that the receptor can be transported to cilia for sensing environmental cues.

10.1126/science.1157983

CLIMATE CHANGE

Evidence for Upwelling of Corrosive "Acidified" Water onto the Continental Shelf

R. A. Feely et al.

As a result of anthropogenic CO_2 uptake, corrosive seawater undersaturated with calcium carbonate shoaled on the continental shelf of western North America in 2007.

10.1126/science.1155676

LETTERS

TECHNICAL COMMENT ABSTRACTS

PHYSIOLOGY

Comment on "Brain IRS2 Signaling Coordinates Life Span and Nutrient Homeostasis" 1012
C. Selman, S. Lingard, D. Gems, L. Partridge, D. J. Withers
full text at www.sciencemag.org/cgi/content/full/320/5879/1012b

Response to Comment on "Brain IRS2 Signaling Coordinates Life Span and Nutrient Homeostasis"
A. Taguchi and M. F. White
full text at www.sciencemag.org/cgi/content/full/320/5879/1012c

CREDIT: (BOTTOM IMAGE) NASA/JPL/CORNELL UNIVERSITY

BOOKS *ET AL.*

POLICY FORUM

PERSPECTIVES

BREVIA

MICROBIOLOGY

CONTENTS continued >>

RESEARCH ARTICLES

REPORTS

AAAS
ADVANCING SCIENCE, SERVING SOCIETY

SCIENCE (ISSN 0036-8075) **is published weekly on Friday, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue, NW, Washington, DC 20005.** Periodicals Mail postage (publication No. 484460) paid at Washington, DC, and additional mailing offices. Copyright © 2008 by the American Association for the Advancement of Science. The title SCIENCE is a registered trademark of the AAAS. Domestic individual membership and subscription (51 issues): $144 ($74 allocated to subscription). Domestic institutional subscription (51 issues): $770; Foreign postage extra: Mexico, Caribbean (surface mail) $55; other countries (air assist delivery) $85. First class, airmail, student, and emeritus rates on request. Canadian rates with GST available upon request, GST #125439322. Publications Mail Agreement Number 1069624. **Printed in the U.S.A.**

Change of address: Allow 4 weeks, giving old and new addresses and 8-digit account number. **Postmaster:** Send change of address to AAAS, P.O. Box 96178, Washington, DC 20090–6178. **Single-copy sales:** $10.00 current issue, $15.00 back issue prepaid includes surface postage; bulk rates on request. **Authorization to photocopy** material for internal or personal use under circumstances not falling within the fair use provisions of the Copyright Act is granted by AAAS to libraries and other users registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided that $20.00 per article is paid directly to CCC, 222 Rosewood Drive, Danvers, MA 01923. The identification code for Science is 0036-8075. Science is indexed in the *Reader's Guide to Periodical Literature* and in several specialized indexes.

CONTENTS continued >>

SCIENCE**NOW**

www.sciencenow.org
HIGHLIGHTS FROM OUR DAILY NEWS COVERAGE

Monkey Model of Huntington's Disease
Genetically modified primates may be better than mice for studying neurological disorders.

Astronomers in a Spin About Mystery Pulsar
One of the universe's most extreme objects just got a bit stranger.

Catching a Climate Offender
New strategy could reduce CO_2 emissions from coal plants.

An unhappy microbe-host interaction.

SPECIAL SECTION

Microbial Ecology

SCIENCE **SIGNALING**

www.sciencesignaling.org
THE SIGNAL TRANSDUCTION KNOWLEDGE ENVIRONMENT

EDITORIAL GUIDE: Focus Issue—A Niche of One's Own
E. M. Adler and J. F. Foley
The nature of microbe-host relationships often depends on signaling pathways in the host.

PERSPECTIVE: Diversification of the Function of Cell-to-Cell Signaling in Regulation of Virulence Within Plant Pathogenic Xanthomonads
M. Dow
Different plant pathogens use similar signaling molecules in distinct ways.

PERSPECTIVE: Bacterial-Modulated Signaling Pathways in Gut Homeostasis
W.-J. Lee
Stimulation of the production of reactive oxygen species in gut epithelial cells by commensal bacteria dampens the host immune response.

PERSPECTIVE: ETosis—A Novel Cell Death Pathway
F. Wartha and B. Henriques-Normark
Pathogenic microbes are trapped and killed by mast cell– and neutrophil-derived extracellular traps.

Telling the story of teamwork.

SCIENCE **CAREERS**

www.sciencecareers.org/career_development
FREE CAREER RESOURCES FOR SCIENTISTS

MiSciNet: Family Trailblazers
S. Gaidos
Fitting in on a college faculty is harder when you are the first in your family to go to college.

Tooling Up: Transitioning to Teamwork
D. Jensen
How do you convince a recruiter that you can play well with others?

Mastering Your Ph.D.: Careers in Management Consulting
B. Noordam and P. Gosling
Problem-solving and communication are important skills if you want to advise industry executives.

From the Archives: How to Get a Job in Academia
A. Fazekas
Search committees at different institutions look for different strengths in their faculty applicants.

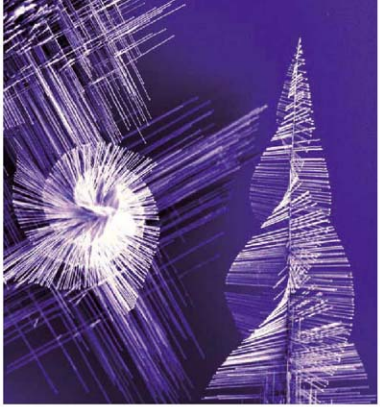
SCIENCE**PODCAST**

www.sciencemag.org/about/podcast.dtl
FREE WEEKLY SHOW

Download the 23 May *Science* Podcast to hear about silica deposits on Mars, a food-driven circadian clock in mice, creating a lake in Turkmenistan, and more.

Separate individual or institutional subscriptions to these products may be required for full-text access.

CREDITS: (SCIENCE CAREERS) COMSTOCK; (SCIENCE SIGNALING) FRITZ WESTOVER/TEXAS A&M UNIVERSITY/TEXAS COOPERATIVE EXTENSION



CREDITS: (TOP TO BOTTOM) M. J. BIERMAN, Y. K. A. LAU, S. JIN; LAMB *ET AL.*

<< Building Nano-Pine Trees

While nanoscale wires and rods hold tremendous potential in electronics, optics, or catalysis, one limitation is in the patterning or development of hierarchical structures. Branched wires can be grown using metal nanoparticles or by changing the composition or growth conditions of the wire, but **Bierman *et al.*** (p. 1060, published online 1 May) now show a growth mechanism that is driven solely by the screw component of a single dislocation along the long axis of the nanowire. The nanowires have a pine tree morphology, with regular branches spawning off a thicker trunk, and the growth agrees with the theory of how screw dislocations relax at surfaces.

Improving the Strength of Steel

Steel will typically lose strength and ductility at lower temperatures, becoming weaker and more brittle, which may have contributed to the sinking of the Titanic. Solving this problem usually requires the addition of a number of alloying elements that can significantly raise the cost of the steel. **Kimura *et al.*** (p. 1057; see the Perspective by **Morris**) have developed a low-alloy steel with elongated grain structure that shows an inverse temperature dependence of the strength and ductility, so that the materials' properties improve as the temperature is lowered.

Regulating S-Nitrosylation

Covalent modification of proteins by S-nitrosylation is an important mechanism for regulation of biochemical activity in cells. However, mechanisms of protein denitrosylation have not been well characterized. The protease caspase-3, which promotes apoptosis, is inhibited by S-nitrosylation and is denitrosylated in cells in which the cell death–promoting receptor Fas is activated. **Benhar *et al.*** (p. 1050; see the Perspective by **Holmgren**) purified a protein fraction that catalyzed denitrosylation of caspase-3 and identified thioredoxin-1 (Trx1) as the protein most likely to be responsible for the denitrosylation activity. Depletion of Trx1 caused accumulation of S-nitrosylated caspase-3 and other S-nitrosylated proteins in cultured cells, and Fas-induced denitrosylation of caspase-3 was inhibited by depleting thioredoxin reductase 2. Thus, regulated denitrosylation of target proteins by Trx1 appears to provide a key component of enzymatic regulation of caspase-3 and possibly other proteins by S-nitrosylation.

Here Comes the Flood

One piece of evidence supporting the presence of past groundwater on Mars has been canyons that start in a steep amphitheater shape. Similar canyons on Earth have been interpreted as marking areas of seepage and flow of groundwater.

Lamb *et al.* (p. 1067) now show that one of the examples, Box Canyon, Idaho, on the Snake River Plain, is likely to have formed during a glacial-aged megaflood, not by gradual erosion caused by groundwater seepage. Cosmogenic dating of boulders in the canyon suggests that the flood was about 45,000 years ago.

Sodium Gets Complex Under Pressure

Until recently, the pressure-temperature-phase diagram of elemental sodium, while exhibiting a curious minimum melting transition under high pressure, was thought to be understood fully. Using state-of-the-art, high-pressure x-ray synchrotron single-crystal diffraction techniques to focus in on the minimum melt regime, **Gregoryanz *et al.*** (p. 1054) report a remarkable concentration of different crystalline phases (6 within a 4-GPa window) that were not predicted by theory. While the mechanism giving rise to such complex behavior remains unclear, such structural complexity may be par for the course, even for simple elements like hydrogen.

Fishing for Viruses

We understand very little about how microbial viruses influence the natural environment. Using sequence data gathered from acid mine microbial communities, **Andersson and Banfield** (p. 1047) have taken advantage of the recent discovery of what appears to be a prokaryotic defense mechanism against viruses to identify "wild" viruses in uncultured biofilms. Several classes of viruses were identified, and the analysis suggests that a rapidly changing population of viruses and host microbes is engaged in a defensive arms race of sequence shifting.

Martian Opal

Extensive deposits of nearly pure silica—either as quartz or opal—are indicative of complex secondary processes, including the involvement of fluids. **Squyres *et al.*** (p. 1063) now report the discovery of opaline silica deposits on Mars by the rover Spirit. Thermal emission data show that the deposits, which include nodules on an outcrop, rock samples, and light-colored soil excavated by the rover wheel, are opal, not quartz or cristobalite. Formation is likely to have involved hydrothermal fluids at low pH, perhaps associated with past volcanic activity in the region.

Continued on page 983

Continued from page 981

Two Clocks Are Better Than One

The circadian clock of mammals resides within the hypothalamus, in the suprachiasmatic nucleus, and it is entrained by light. There is evidence that another circadian clock—entrained by food—also exists in the brain, possibly elsewhere in the hypothalamus. By replacing a missing clock component (*Bmal*) in selected hypothalamic nuclei of mice lacking *Bmal1*, **Fuller *et al.*** (p. 1074) identified the dorsomedial nucleus of the hypothalamus as the site of the food-entrained clock. When *Bmal1* was reintroduced selectively into the suprachiasmatic nucleus, the previously unresponsive mice regained the ability to entrain both locomotor activity and body temperature rhythms to a 12:12 light:dark cycle. Reintroduction of *Bmal* into the dorsomedial nucleus restored the ability of the animals to entrain to a restricted period of food availability, but not to a light cycle. Thus, the dorsomedial nucleus contains a second circadian clock that seems to be induced when food is restricted to take over control of functions such as activity levels that increase the chances of the animal successfully locating additional food resources.

Widening RNA Interference Functions

In plants, fungi, and worms, RNA interference (RNAi) functions to silence endogenous repeated and parasitic DNA, including transposons. In flies and mammals, a distinct RNA-based silencing pathway involving Piwi-interacting RNAs (piRNAs) operates in the germ line to silence the parasitic DNA there. **Ghildiyal *et al.*** (p. 1077, published online 10 April; see the Perspective by **Birchler and Kavi**, published online 8 May) now show that *Drosophila* somatic cells produce endogenous small interfering RNAs (siRNAs). Like piRNAs, many of the endogenous siRNAs map to large genomic clusters that include transposons and repeated sequences, but also to messenger RNAs (mRNAs), suggesting that endogenous siRNAs may regulate mRNA expression.

A Sweet Tale of Gene Hunting

Epidemiological studies have shown that even modest elevations in fasting plasma glucose (FPG) levels predict an elevated risk of cardiovascular disease, even when these levels fall within the "normal" range and are not associated with type 2 diabetes. The molecular mechanisms controlling FPG levels in humans are incompletely understood. Using data from a genome-wide association study, **Bouatia-Naji *et al.*** (p. 1085) identified a single nucleotide polymorphism (SNP) that contributes to the interindividual variation in FPG levels in normoglycemic individuals. The SNP resides within an intron of the *G6PC2* gene, which encodes glucose-6-phosphatase catalytic subunit-related protein (also known as IGRP). This protein is selectively expressed in pancreatic islets and is thought to function by modulating the set point for glucose-stimulated insulin secretion.

Unmasking the Sensor of Iron Deficiency

Iron deficiency is common among animals. Human cells can sense and utilize iron taken up by the body from food sources; however, the mechanism for perception of low iron availability has remained obscure. By positionally cloning the gene mutation responsible for the mouse "mask" phenotype, characterized by iron deficiency anemia and progressive loss of body (but not facial) hair, **Du *et al.*** (p. 1088, published online 1 May) have identified a critical element of the "low iron" detection apparatus. The cell surface protease transmembrane serine protease 6 acts to inhibit the production of hepcidin, the master repressor of intestinal iron absorption. The transmembrane serine protease 6 mutation involved in the mask phenotype specifically impairs dietary iron absorption, highlighting the importance of this homeostatic pathway.

CREDIT: DU *ET AL.*

Kindness or Fairness?

It's easy to propose that allocations of scarce resources should provide the greatest benefit to a group as a whole and be as fair as possible to individual members of the group, but what should be done when both aims cannot be optimized simultaneously? **Hsu *et al.*** (p. 1092, published online 8 May; see the 9 May news story by **Miller**) use functional brain imaging, not to resolve these dilemmas, but to probe the underlying cognitive and emotional processes supporting one view (favoring equity, for instance) versus the other (maximizing the good). Brain regions involved in encoding reward relate also to calculations of total benefit, whereas the balancing of equity and utility seems to be the province of the insula, which connects with emotion-processing neural systems. Thus, judgments of fairness derive from emotion-based preferences, rather than those of pure reason.

James Tiedje is a professor of microbiology and crop and soil sciences and director of the Center for Microbial Ecology at Michigan State University, East Lansing, MI.

Timothy Donohue is a professor of bacteriology at the University of Wisconsin, Madison, WI, and principal investigator of the Great Lakes Bioenergy Research Center, Madison, WI.

Microbes in the Energy Grid

THE CURRENT SURGE IN FOOD AND FUEL PRICES HAS SOUNDED AN ALARM SHOWING WHY providing a sustainable global energy supply and minimizing climate change are arguably two of the greatest challenges facing 21st-century society. With adequate research and proper implementation, the diverse and often unseen inhabitants of the microbial world—bacteria, yeasts, fungi, and archaea—can help address these challenges.

The incredible metabolic diversity of today's microbial world reflects the accumulated evolutionary response to diverse environments present over the 3.5 billion years that they have inhabited Earth. Over this time, extensive mutation, recombination, and lateral gene transfer have produced microbes that access most of the theoretically possible redox reactions, capture solar and other energy resources, and inhabit a wide range of environments. Some of these microbes—the extremophiles—exist at the current limits of life.

To harness microbial activities to address these challenges, much remains to be learned about the chemical space they occupy. Of the estimated 10^{30} microbes on Earth, the vast majority are still unknown. Unknown microbes live on and in us; some we encounter regularly in food, water, or air, and others live in often inhospitable places. Most microbes are beneficial, and their combined activities positively affect numerous aspects of the biosphere and its inhabitants.

For energy production, microbes offer efficient and sustainable ways to convert plants or other biomass into liquid fuels, hydrogen, methane, electricity, or chemical feedstocks currently derived from fossil fuels. But to realize this potential, society must invest in research programs to decouple bioenergy production from the food chain. Currently, microbial fermentation of sugars and starch from food crops is the main source of the ethanol blended with liquid transportation fuels. However, there is a large underutilized resource of cellulosic biomass from trees, grasses, and the nonedible parts of crops that could serve as a feedstock. To tap into this cellulose energy reserve, public and private sectors have begun research programs to use metagenomic, synthetic, and other approaches to identify microbes, enzymes, or microbial communities that release sugars from cellulose and convert them into ethanol or other fuels. Once efficient technologies are capable of achieving this, large-scale conversion of cellulosic biomass into ethanol and other biofuels, including hydrocarbons, will reduce the need to produce fuel from edible sugar and starch.

Microbial-plant relationships can improve the sustainability of biofuel production. Microbes surround plant roots; inhabit stems, roots, and leaves; and live on leaf and root surfaces. Individual microbial activities also provide plants with nitrogen (decreasing the need for fertilizer), help them access phosphorus, protect them from disease, recycle nutrients, and improve soil structure. These associations can improve crop production, especially on marginal lands, and benefit both food and fuel production, helping to abate the current food-versus-fuel debate and lessen the environmental footprint of agriculture.

Plants and autotrophic microbes might also help mitigate climate change because they sequester atmospheric carbon dioxide at a rate that dwarfs its flux into fossil fuels (120 versus 7.6 petagrams of carbon per year). It's the subsequent oxidation of soil and plant carbon by microbes that determines how much carbon will remain sequestered. Thus, managing the dynamics of the plant-microbe-soil ecosystem warrants further investigation, as it could improve the productivity and carbon neutrality of cellulose-to-liquid fuel systems, while helping us reshape how society generates a substantial fraction of its energy.

As Earth's master chemists, microbes can help address society's energy, environmental, and food challenges. To realize these goals, the scientific community must inform the public and policy-makers about the research needed to bring the chemical and catalytic power of microbes to bear on meeting our ever-growing energy needs. With proper research coordination and implementation, the genetic blueprints of microbes will provide a rich stockpile of building blocks to provide new energy sources while improving the health of the biosphere.

– James Tiedje and Timothy Donohue

10.1126/science.1159999

ECOLOGY

A Steady Supply of Food

The environmental and anthropogenic factors affecting leatherback turtle populations are generally unknown and cannot explain why populations are increasing in the Atlantic yet declining in the Pacific. By analyzing reproductive success rates, migratory patterns, and oceanographic variables, Saba *et al.* demonstrate that turtle populations are most likely regulated by the abundance of zooplankton where they forage, as measured by the net primary productivity (NPP) estimated from satellite data. NPP was shown to vary among oceanic regions and to be affected by regime shifts that typically enhance ocean productivity; however, both the total NPP and overall consistency in NPP over time were much lower in the Pacific regions examined than in the Atlantic. In turn, these differences in local resources appear to affect turtle size and overall egg production—larger females that have matured faster lay more eggs than smaller females. These differences, coupled with anthropogenic impact and the effects of climatic shifts on NPP, may explain the discrepancy in demography between the Atlantic and Pacific leatherback populations. — LMZ

Ecology **89**, 1414 (2008).

APPLIED PHYSICS

Tracking Gyrating Memories

Whereas the bits in present magnetic memory storage media are written using electrical current pulses in a write head, recent work has shown that the spin-polarized current in a ferromagnetic medium can impart a spin-transfer torque to the magnetization, resulting in its motion. Such a wired-up memory architecture would allow yet higher storage density and provide faster access times than present hard drives offer. Understanding how the spin current affects the dynamics will help in developing memory storage devices based on this effect. Bolte *et al.* provide a direct view of the dynamics of the process. They use time-resolved x-ray imaging to track the motion of the magnetization, showing how magnetic vortices in permalloy dots gyrate and jiggle in response to the injected spin-polarized current. The technique is sufficiently sensitive to uncover the various contributions to the dynamics, thereby offering the opportunity to fine-tune the process. — ISO

Phys. Rev. Lett. **100**, 176601 (2008).

CHEMISTRY

Does Saltwater Wobble?

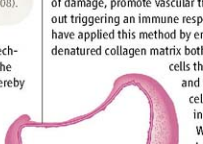
When ionic salts dissolve in water, the resultant solution tends to become more viscous than the pure liquid. This observation can be broadly understood based on a picture in which the web of hydrogen bonds holding the water molecules together rigidifies around solvated ions. However, the molecular details appear much more complicated, because time-resolved vibrational spectroscopy has suggested that most individual water molecules continue to rotate freely regardless of how much dissolved salt is present. Turton *et al.* explore this discrepancy using two related spectroscopic techniques to probe the molecular structure of aqueous salt solutions. The first, dielectric relaxation spectroscopy, is sensitive to individual molecules' orientations, and confirms the rotational freedom previously observed. The second technique, optical Kerr effect (OKE) spectroscopy, reflects polarizability and so is more sensitive to intermolecular changes that stem from translation. The OKE data reveal increasing translational restriction with rising salt concentration, as the ion solvation shells crowd against one another. The authors note that such a decoupling of rotation and translation is analogous to the jamming that occurs during transitions from a liquid to a glass upon supercooling. — JSY

J. Chem. Phys. **128**, 161102 (2008).

BIOMEDICINE

Engineering a Healing Environment

In respiratory conditions such as asthma and chronic obstructive pulmonary disease, the airways narrow, leading to impaired oxygen exchange. Tissue engineering affords one approach to reverse such damage: Endothelial cells embedded in a polymer matrix adopt normal morphologies and, when implanted close to sites of damage, promote vascular tissue repair without triggering an immune response. Zani *et al.* have applied this method by embedding into a denatured collagen matrix both the epithelial cells that line airways and the endothelial cells from surrounding tissue layers. Wrapping injured trachea in this cell-containing matrix enhanced recovery as measured by the size of the lumen and extent of the injury. Either cell type alone had beneficial effects, but both together synergized to improve luminal size and epithelial area more effectively. Measurements of cytokine and growth factor secretion from the endothelial and epithelial cells



Trachea lumen surrounded by epithelium, mesenchyme, and cartilage.

Continued on page 989

CREDITS (TOP TO BOTTOM): DANIEL RYNECKI/GETTY IMAGES; ADAPTED BY ZANI *ET AL.*, PROC. NATL. ACAD. SCI. U.S.A. 105, 19149 (2008)

Continued from page 987

indicate that the two cell types are interdependent. Their crosstalk increases healing, presumably via secretion of repair-promoting factors, but because the cells are interspersed within the matrix, the normal cellular architecture of the tissue seems not to be required. — KK

Proc. Natl. Acad. Sci. U.S.A. **105**, 7046 (2008).

GEOLOGY

P in All Its Guises

Phosphorus is the rarest of the major biological nutrients on Earth, tending to be locked up primarily in a few rather insoluble minerals. Nonetheless, it plays a key role in biological processes and metabolism, feeds primary production in the oceans and plants on land, and is required to form bones, teeth, and some shells. Rampant soil fertilization and industrial use have conversely led to the problem of too much phosphorus in many rivers and estuaries, where it is difficult to remove.

Filippelli considers the long-term geological cycling of phosphorus and how this critical element gets released to the environment over time or sequestered. Erosion enhances the release of phosphorus, which makes its way to the oceans. Records of past marine phosphorus concentrations and biological productivity suggest that the formation and uplift of the Himalayas may have thus fueled ocean productivity ~20 million years ago; similar arguments have been proposed for glacial cycles. Further in the past, the relation between enhanced erosion and thick sequences of phosphorus marine rocks remains unclear. Today, the phosphorus cycle is dominated by human activities; we have doubled the natural riverine load globally. As discussed by Oelkers and Valsami-Jones, further use at this rate may soon run up against the limited geological supply. — BH

Elements **4**, 89; 83 (2008).

MICROBIOLOGY

A Specialist Repertoire

Sequencing the genome of the fungus *Trichoderma reesei* has led to surprising questions about how it degrades biomass efficiently and whether it can be engineered for the commercial production of biofuels. Breaking down lignocellulose in plant cell walls requires cellulases and hemicellulases, but Martinez *et al.* found that *T. reesei* harbors fewer genes dedicated to cellulose digestion than do 13 other fungi that digest plant cell walls. Some of these genes are located in clusters, which may reflect a functional organization that allows efficient enzyme production. Moreover, genes encoding the secretory machinery of *T. reesei* are similar to those of budding yeast, consistent with its extraordinary ability to secrete 100 g of protein per liter. The absence of a wider variety of cellulose-degrading enzymes may be an opportunity to engineer industrial strains of *T. reesei* for economical processing of biomass feedstocks used in the production of biofuels. — LC

Nat. Biotechnol. **26**, 553 (2008).

Science Signaling
AAAS

<< Lateral Receptor Transfer

Although many malignant gliomas display EGFRvIII, an oncogenic mutant form of the epidermal growth factor receptor (EGFR), the mutant receptor may be expressed only by a subset of cells in the tumor. Al-Nedawi *et al.* found that a human glioma cell line (U373) transfected with EGFRvIII (U373vIII) produced more membrane-derived microvesicles than did the parent cell line. The U373vIII cell-derived microvesicles contained EGFRvIII, as well as the lipid raft marker flotillin-1. Unlike U373 cells, U373vIII cells formed subcutaneous tumors when injected into immunodeficient mice, and these tumors released EGFRvIII-containing microvesicles into the circulatory system. U373 cells that had been exposed to EGFRvIII-containing microvesicles showed increased phosphorylation of extracellular signal–regulated kinase 1 and 2 (ERK1/2). Pharmacological treatment to inhibit signaling by EGFR-family receptors attenuated this increase in ERK1/2 phosphorylation, as did treatment with annexin V (to prevent microvesicle uptake). EGFRvIII-containing microvesicles also stimulated other downstream events: release of vascular endothelial growth factor, levels of the anti-apoptotic protein Bcl-x_L, and the ability of U373 cells to grow in soft agar (indicative of malignant transformation). The authors conclude that membrane microvesicles provide a pathway for transfer of the transformed phenotype. — EMA

Nat. Cell Biol. **10**, 619 (2008).

CREDIT: NEW YORK STATE MUSEUM

www.sciencemag.org

Science

1200 New York Avenue, NW
Washington, DC 20005
Editorial: 202-326-6550, FAX 202-289-7562
News: 202-326-6581, FAX 202-371-9227
Bateman House, 82-88 Hills Road
Cambridge, UK CB2 1LQ
+44 (0) 1223 326500, FAX +44 (0) 1223 326501

Published by the American Association for the Advancement of Science (AAAS), *Science* serves its readers as a forum for the presentation and discussion of important issues related to the advancement of science, including the presentation of minority or conflicting points of view, rather than by publishing only material on which a consensus has been reached. Accordingly, all articles published in *Science*—including editorials, news and comment, and book reviews—are signed and reflect the individual views of the authors and not official points of view adopted by AAAS or the institutions with which the authors are affiliated.

AAAS was founded in 1848 and incorporated in 1874. Its mission is to advance science and innovation throughout the world for the benefit of all people. The goals of the association are to: foster communication among scientists, engineers and the public; enhance international cooperation in science and its applications; promote the responsible conduct and use of science and technology; foster education in science and technology for everyone; enhance the science and technology workforce and infrastructure; increase public understanding and appreciation of science and technology; and strengthen support for the science and technology enterprise.

Information for Authors

See pages 634 and 635 of the 1 February 2008 issue or access www.sciencemag.org/about/authors

Sculpture Diagnosis

Erected in 1980 but dismantled 16 years later because of concerns about its stability, Henry Moore's sculpture *Large Arch* may soon be back in place in London's Kensington Gardens. The 6-meter-tall work is made of seven pieces of travertine stone connected to resemble sheep collarbones. Experts at Imperial College London, the Glasgow School of Art, and the Tate museum used laser scanning and computer modeling to find the sculpture's weak points. (The most stressed areas are shown in blue.) With the insertion of fiberglass rods and a new base of reinforced concrete, the sculpture could be safe for the ages, the team says—providing that funding can be raised for the job.

Hold 'em or Fold 'em

Competitive wrestling meets molecular biology in Foldit, an online game that lets you tussle with the stuff of life, folding strings of amino acids into their correct protein structures.

Predicting the 3D structure of proteins from amino acid chains takes supercomputers trillions of calculations. In recent years, the public has been helping via data-crunching screensaver programs. Foldit asks people to add the idle processing of their own brains.

After teaching how different amino acids constrain folding, the game presents players with strings of amino acids. The goal is to stabilize a protein's active site by hydrogen-bonding strands and reducing its total volume—but not by too much, because bringing like-charged amino acids together results in a bristle of red alarms. The quicker you get the correct fold, the higher your score.

"My dream is that a 12-year-old in Indonesia will turn out to be a prodigy and build a cure for HIV," says Foldit team leader David Baker, a molecular biologist at the University of Washington, Seattle, in a press release. Arthur Olson, a molecular biologist at the Scripps Research Institute in San Diego, California, doubts that players will perform better than automated computer programs, but he admires Foldit's educational potential, calling it "a puzzle that makes Sudoku look like tic-tac-toe." www.fold.it

High Eye-Q

Do glasses make you look brainy? Apparently, even small children think so.

Eye doctor Jeffrey Walline of Ohio State University in Columbus headed a study to see what children think of other children who wear glasses. Eighty children ranging in age from 6 to 10 were shown pictures of 24 pairs of children, one with and one without eyeglasses, of both sexes and varying ethnicities. They were asked questions including which member of the pair looked smarter.

Which girl looks smarter to you?

The result, published in the May issue of the journal *Ophthalmic and Physiological Optics*, indicates that although gender has a strong influence on some impressions—boys were perceived as better athletes regardless of whether they wore glasses, for example—glasses trump other characteristics when it comes to brains. Female glasses-wearers, in particular, had a 72% likelihood of being seen as smarter than the nonbespectacled children they were paired with. Ethnicity had no effect on choices.

There's a "well-known correlation between [high] IQ and myopia," says behavioral geneticist Nicholas Martin of the Queensland Institute of Medical Research in Brisbane, Australia. Were the children picking up on that? Walline doubts it. "Myopia typically doesn't develop until age 8 or later," so they wouldn't see much of it in their peers, he says. But it does seem that even when you're in first grade, "spectacles make children appear to be smarter."

RELIVING IRAQ

A Humvee heads up a desert road in Virtual Iraq, an emerging treatment for veterans with post-traumatic stress disorder. At this month's meeting of the American Psychiatric Association in Washington, D.C., psychologist Barbara Rothbaum of Emory University in Atlanta, Georgia, reported promising results for a technique that combines Virtual Iraq with a drug that modifies the brain's fear response. The drug, D-cycloserine, enhances the function of a receptor for the neurotransmitter glutamate—the so-called NMDA receptor—that is critical for memory extinction. Earlier research showed that it helped people reduce their fear of heights (*Science*, 2 April 2004, p. 34).

In each of five sessions, soldiers take the drug and don virtual-reality goggles. Then a therapist guides them through a traumatic memory, most often an encounter with an improvised explosive device. The experience comes with sounds—people yelling, dogs barking, guns discharging, and helicopters whirring—vibrations, and even smells of burning rubber and fuel. "In general, veterans don't respond as well as civilians to drugs or therapy," Rothbaum said, but this combination makes for a "more potent exposure." The researchers have so far enrolled 27 vets, with 1-year follow-ups on three patients. Preliminary data, she said, indicate that two sessions with the drug achieve as much as eight without it.

CREDITS: (TOP TO BOTTOM) THE HENRY MOORE FOUNDATION; JEFFREY WALLINE, SKIP RIZZO

IN THE COURTS

PROLIFIC. An Istanbul court has sentenced an influential Islamic creationist to 3 years in prison for starting a criminal organization and profiting from it. But the conviction, which Adnan Oktar says he will appeal, seems unlikely to stem the flood of creationist books and DVDs he is publishing.

Oktar, who uses the pen name Harun Yahya, became well-known outside Turkey when his Foundation for Scientific Research (BAV) widely distributed its *Atlas of Creation*, a stunning, 768-page tome (*Science*, 16 February 2007, p. 925). BAV is not directly linked to the activities that landed Oktar in trouble, and creationism had nothing to do with the charges.

Even so, a BAV spokesperson says Oktar is being persecuted "because of his ideas." Given the political pressures on Turkey's justice system, that's "not entirely implausible," says physicist Taner Edis of Truman State University in Kirksville, Missouri, who has followed the case closely. BAV says "the work will go on" even if its leader goes to prison.

RISING STARS

FOOT SOLDIERS. At age 13, Anya Suslova thought it would be fun to tag along with her father, a boat captain in Zhigansk, Siberia, as he helped an international team of climate scientists collect Arctic water samples along the Lena River. Her curiosity and hard work—she volunteered to collect samples during the bitter Russian winter, including drilling through 0.6 m of ice—have spawned the Student Partners Project, a science and outreach effort funded by the U.S. National Science Foundation (NSF).

Anya's effort in 2003 "was the start of a fantastic addition" to the project's data set, says team leader R. Max Holmes, an earth systems scientist at Woods Hole Research Center in Massachusetts. In 2005, Holmes and his colleagues received a $630,000 NSF grant to study the effects of climate change on Arctic river systems with help from hundreds of schoolchildren, their teachers, and community leaders.

Suslova, now studying economics and political science at Yakutsk State University, is spending this summer at Woods Hole and plans to pursue a career in environmental policy. "Zhigansk is developing. I want it to be environmental[ly] clean [so] that we will be able to drink water from Lena as we do now and swim there."

CREDITS: (TOP TO BOTTOM) PATRICK ROBERT/CORBIS; SOURCE: GARY MITTLEMAN; MICROSOFT; MOLLY INGELBOY

PIONEERS

AMONG THE STARS. Jim Gray went sailing off the coast of San Francisco in January 2007 and was never seen again. But last week, Microsoft, his employer, unveiled one of his legacies: a way to help astronomy buffs explore the night sky from their computers.

WorldWide Telescope is a program that brings together high-resolution images from several ground-based and space-based telescopes. It builds upon an online astronomical image database called the SkyServer that Gray pioneered back in 2001. The application enables users to "fly" through a virtual sky, zooming in on constellations, galaxies, and stars of interest, and even take "tours" guided by astronomers.

"Aside from being a brilliant database engineer, he was a great connector," says Curtis Wong, a computer visualization expert at Microsoft Research in Redmond, Washington, who led the effort to create worldwidetelescope.org. Wong says that Gray, who worked at Microsoft for 11 years, helped him team up with astronomers and other researchers from 10 institutions around the world. Wong's team has dedicated its project to their former colleague.

Nonprofit World

BOTTOM-UP. Fuel-cell developer and millionaire Gary Mittleman decided to run for Congress to push for better federal policies on energy, the environment, and the economy. But last month, the 55-year-old Democrat changed his mind after deciding that he could get more bang for his bucks—he had planned to donate up to $300,000 to his campaign to represent upstate New York in the U.S. House of Representatives—by setting up a nonprofit organization.

Mittleman says that the One Dream One Earth Foundation, which he is seeding with funds that he and his donors plan to provide, will conduct an education campaign at the grass-roots and federal levels. He hopes that better information will dissuade lawmakers from taking steps such as starving the U.S. strategic petroleum reserve of oil, an idea that he says is simply "more lunacy."

Got a tip for this page? E-mail people@aaas.org

A flood of lawsuits
1000

The skinny on microbes
1001

On the fault. A massive landslide crushed some buildings in Beichuan.

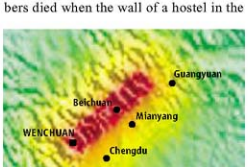
SICHUAN DISASTER

Landslides, Flooding Pose Threats As Experts Survey Quake's Impact

CHENGDU, CHINA—Wei Fangqiang knows what it's like when a mountain crumbles: The Longmenshan, or Dragon's Gate Mountains, are prone to landslides. But when the physical geographer and seven colleagues with the Chinese Academy of Sciences' Institute of Mountain Hazards and Environment (IMHE) in Chengdu trekked into the area devastated by the Sichuan earthquake, they were stunned. It looked as though the hills had been blown apart. Landslides had flattened several-story buildings in the town of Beichuan and annihilated villages that clung to the steep slopes. In Wenchuan, Wei and his comrades picked their way across a 70-meter-high, 300-meter-wide rubble pile that had crushed a hydropower station and blocked the Chaping River. If an aftershock had struck, it could have spawned a new landslide where they were walking. "It was very, very dangerous," Wei says.

Landslides unleashed by the rupture of a more than 200-kilometer section of the Longmenshan fault, followed by powerful aftershocks, dammed parts of nine rivers, creating 24 new lakes. The biggest and most threatening is 3.5 kilometers upstream of Beichuan. If the debris dam were to break, the resulting flood would threaten relief workers and researchers in Beichuan. "We're worried about another catastrophe," says Wei. As *Science* went to press, experts with the Ministry of Water Conservation were weighing options for how to relieve pressure building up behind the dam. They had at most a week to act, said Cheng Genwei, IMHE's vice director.

Down the road from IMHE, researchers with the Chengdu Institute of Biology (CIB) were in mourning. Three senior staff members died when the wall of a hostel in the mountains collapsed as they were dashing out of the door for safety. (IMHE lost one staffer in Beichuan.) After a 20 May memorial service, CIB scientists were hoping to return to work with an ambitious research agenda, including an examination of habitat fragmentation and ecological succession in landslide areas. "The earthquake will be a big driver for research," says CIB ecologist Bao Weikai. He and colleagues will also be alert to a grave threat to Sichuan's famed giant pandas: the possibility of a massive die-off of bamboo, the panda's staple, like one recorded in a quake 30 years ago.



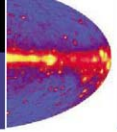
Shake map. The magnitude-7.9 earthquake centered in Wenchuan brought devastation to the severe shock zone (red) on the Longmenshan fault.

At 2:28 p.m. local time on 12 May, the Sichuan earthquake struck with a magnitude of 7.9. It "was not a total surprise to geophysicists," says Mian Liu, a geophysicist at the University of Missouri, Columbia. It occurred on a well-known, active fault system, he notes, which in 1933 produced a magnitude-7.5 quake that killed about 9000 people.

But the death toll of the Sichuan earthquake is horrific. As of 20 May, more than 40,000 people are known to have perished, including thousands of children. Experts are asking whether better construction, especially at schools, could have prevented many deaths. "Earthquakes themselves do not kill people," says Liu. The biggest killer, he says, is structural collapse—"a point so sadly illustrated by this earthquake." It appears that many wrecked buildings were not reinforced. "One hardly sees steel

CREDITS: TOP TO BOTTOM: COURTESY OF WEI FANGQIANG/IMHE; USGS

Five-way symbiosis
1006



Gamma ray vision
1008

beams extruding from the collapsed buildings," Liu says. "When they are seen, they are so thin that they bent with the debris like overcooked noodles."

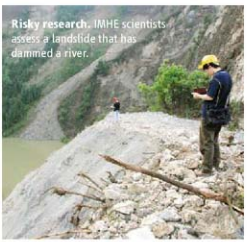
Under a makeshift canopy next to a swimming pool at a community center in the hard-hit historic town of Dujiangyan, west of the epicenter, geophysicist Miao Chong-Gang points to a map on his laptop overlain with seven circles in a line on the Longmenshan fault. It's the latest data from China's seismic monitoring network showing that the Sichuan earthquake was composed of seven powerful sequential ruptures unleashed when the fault ruptured southwest to northeast. "Several years ago, we could not do an analysis like this," says Miao. But with more than 1000 seismometers now in a digital network, China can now parse data like this in a few hours.

Within 30 minutes after the quake hit, the China Earthquake Administration (CEA) in Beijing had crunched the numbers and issued a preliminary forecast of at least 7000 deaths. Their assessment would prove to be an underestimate, but it was alarming enough to prompt CEA to mount a full-scale response. Miao, vice-director of CEA's Earthquake Emergency Management Department Response Command Center, led a 230-person team to Dujiangyan late in the evening on 12 May. His group, one of 187 rescue teams in the disaster area, has saved 48 people; in the morning of 19 May, they were elated to have saved a 61-year-old woman who had survived 163 hours in the rubble.

Miao's team was about to switch from rescue to recovery. Among their tasks over the next 2 months, Miao says, is to ground-truth the computer-generated data. That will mean conducting seismic, strong-motion, and geologic surveys and running tests on everything from geomagnetism to water chemistry. Such research must wait until the aftershocks have subsided. Several CEA volunteers who were ferrying food and water on foot into the disaster zone were among more than 150 relief workers known to have died in aftershock-induced landslides. The slides also claimed the lives of two Sichuan Earthquake Administration researchers who were measuring crust deformation. "We have almost no experience in responding to an earthquake in a mountainous area," says Miao.

Back in Chengdu, CIB scientists are itching to get out into the field. A week after the quake, 10 of their colleagues were alive but stranded at CIB's Maoxian Mountain Ecosystem Research Station in a pine forest 220 kilometers northwest of Chengdu. The institute had a couple of dozen long-term projects in the disaster area, a biodiversity hot spot that encompasses 22 nature reserves. They'll have to write a new research plan. "The earthquake has dramatically changed the landscape," says CIB ecologist Luo Peng.

One urgent task is to monitor bamboo. The plant flowers once every 70 years or so. Shortly after a powerful earthquake in the 1970s, large swaths of bamboo suddenly flowered and died, says CIB ecologist Pan Kai-Wen. How a quake might trigger flowering is a mystery, but a large-scale die-off, he says, could pose a big threat to China's endangered giant pandas.



Risky research. IMHE scientists assess a landslide that has dammed a river.

To map the landslides, Wei and his IMHE colleagues ventured into the danger zone on 15 May. They had to abandon their car where a landslide had blocked the highway and head toward Beichuan on foot. Traveling in the other direction was a ragged stream of refugees. When the researchers reached Beichuan the next day, they found that although many buildings had collapsed from the shaking, many others were demolished by massive boulders. "In some places, the landslides did more damage than the earthquake," Wei says. "We know the rock is very loose here. But still I was surprised that the landslides were so severe." In a nearby village, a woman was on top of a pancaked building. "She was calling her son's name, trying to wake him up." There was no one else around.

Wei and his colleagues could not get past a blocked mountain pass leading to the biggest landslide, a 2-kilometer-long debris flow that had clogged the Qingjiang River. To ward off a catastrophic breach, Cheng says, the preferred option is to dig a canal that drains the lake gradually. If that's impossible, he says, they'll have to blast the dam and allow a more chaotic release. Sichuan's rainy season starts in late June; if the rains start early, before the problem is dealt with, the situation could be very dangerous, says Wei.

The IMHE researchers plan to head into the field as early as next week to sample landslide material and draw topographic maps. A future task is to advise authorities on a safe place to rebuild Beichuan city. The original site will almost surely be abandoned. "It should be a memorial to the earthquake victims—and a reserve for seismic research," says Miao. CIB scientists hope to turn the disaster into an opportunity to advise Longmenshan residents about more sustainable livelihoods in the fragile mountain ecosystem. One practice they want to see ended is farming on the steep slopes. Better forest cover could reduce the landslide risk, says Luo: "We need a new strategy of mountain development."

Others say the Sichuan disaster should stimulate China to rethink its entire approach to earthquake research. "In recent decades, geophysicists have spent too much energy and funding on research on deep-earth structure or tectonics," says Zhou Shiyong, a geophysicist at Peking University. He argues that more attention should be devoted to earthquake prediction. "We could find some precursors," he says, such as abnormal patterns in seismic stress or underwater variation before a huge quake occurs. Miao counters that any precursors of the Sichuan quake were minimal. "They could not have given us any warning," he says.

One thing that will surely come under scrutiny is China's construction standards. "More effort should be devoted to earthquake hazards analysis and management, including developing and enforcing proper building codes, especially for schools, hospitals, and other public buildings," Liu says. For thousands of victims in Sichuan, that lesson came too late. **–RICHARD STONE**

With reporting by Chen Xi and Hao Xin.

CREDIT: COURTESY OF WEI FANGQIANG/IMHE

U.S. AGRICULTURE

Farm Bill Gives Agriculture Research A Higher Profile in the Department

Spending on basic agricultural research in the United States could grow significantly thanks to a massive farm bill that Congress approved overwhelmingly last week. The bill also calls for a larger competitive grants program within the U.S. Department of Agriculture (USDA). "We view this as a real win," says Ian Maw, vice president of the National Association of State Universities and Land-Grant Colleges (NASULGC) in Washington, D.C.

The changes are part of the Food, Conservation, and Energy Act of 2008, a 5-year, $307 billion measure that preserves massive subsidies for farmers. It renames the department's major extramural arm, the Cooperative State Research, Education, and Extension Service, as the National Institute for Food and Agriculture (NIFA). The institute is to be headed by a "distinguished scientist" appointed by the president to a 6-year term. The competitive grants portion of the new institute, to be called the Agriculture and Food Research Initiative, will replace the National Research Initiative (NRI). Its budget will be authorized at $700 million a year, $200 million more than the level for NRI, which actually receives only $180 million a year.

Supporters of agricultural research hope these changes will be more than cosmetic. They have long pressed for an entity within USDA analogous to the National Institutes of Health (NIH) or the National Science Foundation. Although Congress rejected a proposal from President George W. Bush to combine the department's intramural and extramural research programs into an office of science (*Science*, 23 February 2007, p. 1073), the legislation seeks to better coordinate USDA's $2 billion research portfolio by requiring an annual "roadmap." It assigns the job to the undersecretary for research, education, and economics, currently Gale Buchanan, former agriculture dean at the University of Georgia.

The arrangement combines the recommendations of a group headed by William Danforth, former chancellor of Washington University in St. Louis, Missouri, and a proposal from NASULGC. Danforth calls the measure "a great breakthrough" but adds that "what will really be necessary will be to build competitive funding." Almost all of NRI's current budget is spent on "formula-driven" research, says Maw, whereas the new bill designates that 60% must go to basic research. It keeps separate the department's intramural arm, the Agricultural Research Service.

The NIFA chief would also oversee $308 million over 5 years for competitive grants in two new areas: organic crops and "specialty crops," otherwise known as fruits and vegetables. This funding becomes an actual spending level unless Congress explicitly decides otherwise.

Bush says that the bill is too generous on agricultural subsidies. But both houses passed it with veto-proof majorities, and it was expected to become law as early as the end of the week. It would take effect in October 2009.

–CONSTANCE HOLDEN

Veggie power. Research on specialty crops, such as artichokes, would get more attention in new farm bill.

RESEARCH FUNDING

Australia's New Science Budget Gets a Mixed Review

CANBERRA, AUSTRALIA—Two of Australia's science agencies are shedding jobs and trimming programs to comply with a new national budget that's both praised and criticized by research leaders. The spending plan announced by the Labor government last week—its first since coming to power in 2007—provides more money for education initiatives, including a $10.5 billion trust fund for higher education infrastructure, but less for two key players, the nation's premier science agency, the Commonwealth Scientific and Industrial Research Organisation (CSIRO), and the Australian Nuclear Science and Technology Organisation (ANSTO). The cuts are troubling, some say, because the government expects to reap a $20.7 billion surplus over Australia's next annual budget cycle, which starts 1 July.

The reduction at CSIRO is "a disappointment," says Chief Executive Geoff Garrett. Combined with a cut announced previously, it will shrink the agency's appropriation through the 2008–09 budget cycle by roughly $15 million, or just over 2%, to $660 million. "Our aim will obviously be to preserve core capability and the science and research activities that we're doing, … but the arithmetic is such that … there will be some staff losses." He says CSIRO will benefit from funding for energy technology, water management, and climate change adaptation. But it will probably have to cut its 6350-strong workforce by about 100. ANSTO, meanwhile, is to lose about 80 of its 1009 staff as it deals with rising costs and a cut of about 2.6% from its $144 million appropriation in the year ending on 30 June 2009.

Australian Academy of Science President Kurt Lambeck welcomed the education investment fund, saying it will "put us on a path to a world-class higher education and research sector." He is also enthusiastic about a promise to create new scholarships and 1000 fellowships for midcareer researchers: "It creates opportunities for retaining people in Australia and attracting overseas researchers at a stage when they are most productive." But climate change research is a different story: An outlay of $2.2 billion over 5 years for global warming R&D, including clean coal and renewable energy projects, "does not reflect the urgency of the problem," he says.

Lambeck worries that a series of reviews into the national innovation system and universities could be "used as an excuse for inaction." Minister for Innovation, Industry, Science and Research Kim Carr could not be reached for comment.

–CHERYL JONES

Cheryl Jones is a writer in Canberra, Australia.

CREDIT: BILL CURTSINGER/NATIONAL GEOGRAPHIC/GETTY IMAGES

GLOBAL WARMING

Hurricanes Won't Go Wild, According to Climate Models

If you put much faith in the world's most sophisticated climate models, there's good news about how hurricanes will react to global warming. Two new model studies project a modest increase or even a decrease in the frequency and intensity of Atlantic tropical cyclones. "The Atlantic isn't going to be swallowed by repeats of the [disastrous] 2005 hurricane season," concludes hurricane researcher Hugh Willoughby of Florida International University in Miami, who did not take part in the work.

But even some of those involved in the studies urge caution in interpreting the results. "I'm much less sanguine about models solving the problem," says Kerry Emanuel, lead author of one of the papers and a hurricane researcher at the Massachusetts Institute of Technology (MIT) in Cambridge. There's still too much messiness beneath the surface of all such studies, he says.

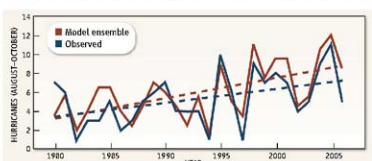
Using two different approaches, both model studies tweak the big global climate models to simulate tropical cyclones. Global climate models can't form tropical cyclones because their picture of the atmosphere is too fuzzy. So, as they report in this week's issue of *Nature Geoscience*, climate modeler Thomas Knutson and colleagues at the National Oceanic and Atmospheric Administration's Geophysical Fluid Dynamics Laboratory (GFDL) in Princeton, New Jersey, put extra computing power into simulating the tropical Atlantic in enough detail to form storms while embedded in a fuzzy global model. And Emanuel and his MIT colleagues randomly "seeded" seven different global models with incipient storms that grew or died depending on whether conditions favored them, as they reported in the March *Bulletin of the American Meteorological Society*.

On average, the two approaches yielded much the same results for the Atlantic, where actual hurricane numbers have doubled in the past 25 years. The GFDL model produced a modest 18% decrease in the frequency of Atlantic hurricanes by the end of this century under global warming and a few percent increase in the intensity of storms. The MIT group reported just a couple of percent increase in frequency and a 7.5% increase in intensity.

Broader interpretations, however, differ. Pointing to their model's striking ability to reproduce variations in hurricane frequency during the past 25 years, Knutson concludes that his group's work "does not support the notion that increasing greenhouse gases will support large increases in hurricane or tropical storm frequency." In contrast, Emanuel finds his "results to be very different when you [run] different models and very different in different ocean basins." For example, using the MIT approach, the GFDL model produces a 23% increase in storm frequency rather than a decrease of 8%, he says. The models do predict a smaller increase in Atlantic hurricane activity than has been seen in the past few decades, Emanuel concludes. That implies that global warming was not the prime driver behind the recent burst of activity.

On the other hand, Emanuel adds, the models may not be properly handling global warming and its effects on tropical cyclones. Other studies have statistically linked the tropical Atlantic warming both to the greenhouse and to the jump in storm activity, many researchers note. What's more, the MIT group's seven different models yield a disturbing variety of predictions—from a 23% increase in frequency to a 29% decrease. And most researchers are concerned that the GFDL group generated input for its Atlantic storm model by averaging together a large range of predicted conditions from global climate models. As might be expected, researchers say bigger and better models are needed to make the message clear.

–RICHARD A. KERR



Hindsight. A Geophysical Fluid Dynamics Laboratory model (red) does well at simulating the actual (blue) year-to-year and long-term variations in hurricane number. The model predicts modest changes under global warming.

CREDIT: ADAPTED FROM T. KNUTSON *ET AL.*, *NATURE GEOSCIENCE*

Paying for Medical School

An Ohio medical school hopes to encourage budding physician-scientists to stick with research by paying for a big chunk of their education. The Cleveland Clinic Lerner College of Medicine, which opened in 2004, accepts 32 students a year into a 5-year program—one more than the norm—specializing in clinical research. Last week, it announced that tuition, worth $43,500 a year, will be free. Current students will get back 50% of what they have already paid.

"We really wanted to remove debt as a potential obstacle to pursuing careers in academic medicine," says spokesperson Raquel Santiago. The scholarships, which are meant to put students on a par with their peers in graduate school, will be financed by endowment funds and clinical income. –JOCELYN KAISER

Help for Mystery Illnesses

The U.S. National Institutes of Health (NIH) has begun a new service for patients with undiagnosed diseases. Some two dozen specialists at its renowned clinical center in Bethesda, Maryland, will be part of a team that will tackle such baffling cases.

These patients often "hit a brick wall," says William Gahl, clinical director of the National Human Genome Research Institute, who will direct the effort. Advances in genetics have led to "more and more … manifestations of new diseases," said NIH Director Elias Zerhouni in announcing the initiative this week.

The new program expects to treat up to 100 patients a year. For more information, call 1-866-444-8806. –JENNIFER COUZIN

A Larger STEM Profile

The U.S. Congress is being asked to elevate the status of science and math education, coordinate $3 billion in current federal programs, and prod states to adopt common K–12 science and math standards. Legislation introduced this week by Representative Michael Honda (D–CA) would create a Science, Technology, Engineering, and Mathematics (STEM) Education Committee within the White House and an Office of STEM Education at the Department of Education. The bill also requires federal STEM programs to share results.

"Having a dedicated office for STEM in the White House should raise its visibility," says Jodi Peterson of the National Science Teachers Association in Arlington, Virginia. A parallel Senate measure was introduced by Democratic presidential hopeful Senator Barack Obama (IL).

–YUDHIJIT BHATTACHARJEE

CLIMATE CHANGE

Polar Bear Listing Opens Door to New Lawsuits

The Bush Administration's decision last week to list the polar bear as a threatened species is about to spark a new round of litigation over greenhouse gas emissions. After analyzing climate models that predict the bear's sea ice habitat would continue to shrink due to global warming, the U.S. Department of Interior ruled that the animal deserves some protection under the Endangered Species Act (ESA). Several environmental groups are preparing to use the ruling to argue that cuts in greenhouse gases are now legally required to protect the polar bear, whereas conservative legal groups are planning to challenge the ruling itself.

When he announced the polar bear's new status, Interior Secretary Dirk Kempthorne tried to preempt litigation to force cuts in greenhouse gases. No specific source of these gases, Kempthorne asserted, will kill any individual polar bear, so the ESA doesn't require power plants, refineries, or even the nation's fleet of automobiles to reduce their emissions. But attorneys on both sides of the long-running legal war over endangered species predict that some courts will reject that argument. "The secretary can't dictate to the courts how they interpret the law," says M. Reed Hopper, a principal attorney for the Pacific Legal Foundation (PLF) in Sacramento, California, a conservative critic of environmental regulation. "I think the environmentalists will find sympathetic judges who will rule that there is a causal connection and give them standing to bring their suits."

Kassie Siegel, an attorney for the Center for Biological Diversity (CBD) in Joshua Tree, California, is leading the environmentalists' strategy. The "attempt to exempt greenhouse gas emissions is illegal and won't stand up," she says. CBD, Greenpeace, and the Natural Resources Defense Council jointly filed their initial legal challenge on 16 May. Siegel, who also filed lawsuits that forced the government to list the polar bear, plans to argue that the ESA requires every government agency to consult with polar bear experts at the U.S. Fish and Wildlife Service before taking any step that could increase emissions of carbon dioxide. Such steps include authorization of oil and gas drilling, issuing permits for coal-fired power plants, or writing new fuel-economy standards for sport utility vehicles and trucks. "It's high time that federal agencies rolled up their sleeves and did what they're supposed to do on greenhouse emissions," she says.

Bearing witness. Lawsuits will use the polar bear's "threatened" status to seek changes in U.S. climate policy.

Some environmentalists doubt that such lawsuits ultimately will reduce greenhouse gas emissions. "I think it's highly unlikely that any court will say, 'This source of emissions has to be halted because it's adding to the burden of carbon dioxide, which is melting ice in Alaska,' " says Michael Bean, a specialist on wildlife conservation at the Environmental Defense Fund in Washington, D.C. Such legal actions could, however, capture public attention, says Holly Doremus, a professor of environmental law at the University of California, Davis. "It can really help people agree that, 'Okay, we've got to act. And we've got to act now.' "

Meanwhile, PLF has announced that it will challenge the Administration's polar bear decision in court. Hopper says it makes no sense for the government to declare the polar bear threatened while insisting that it can do nothing to change the situation. "A listing that cannot address the alleged problem … should not have occurred," says Hopper. "The listing can't affect the melting, but it does open the floodgates to litigation."

Even some environmentalists say the ESA isn't well-designed for dealing with the broad impact of climate change. Doremus points out that global warming may create new dilemmas that the law didn't foresee. "Suppose we decide we can't save

CREDIT: JUPITER IMAGES

The Threat to the World's Plants

A day after polar bears made headlines last week, the world's leading botanical gardens issued a call to remember threatened plants, too. Their new report, *Plants and Climate Change: Which Future?* makes the case for protecting the botanical foundations of terrestrial life. "If you read any report about the impact of climate change, it's almost always about polar bears or tigers," said Suzanne Sharrock, director of Global Programmes for Botanic Gardens Conservation International (BGCI) in London and a co-author of the report.

But BGCI, a network of 2000 organizations involved in plant conservation, says climate change could kill off half of Earth's plant species. Plants that grow on islands or on mountainsides are at greatest risk because they have "nowhere to go" as the climate shifts around them.

BGCI also announced its own global effort to catalog and preserve threatened plants. It will update a 10-year-old survey of the world's trees, identifying species that need additional protection in their native habitat and collecting others for preservation in botanic gardens and arboreta. BGCI plans to reintroduce some threatened plants into their former habitats.

Thomas Lovejoy, president of the H. John Heinz Center for Science, Economics and the Environment in Washington, D.C., welcomed the new initiative. "At the outset, plants were scarcely mentioned in the Endangered Species Act. Now, it's an integral part," he notes. –D.E.C.

all species. Which ones should we concentrate on? The law doesn't allow us to give up easily," she says. "But in situations where we may have to give up on some, we may need better mechanisms for triage." And Bean finds it "worrisome" that "you have a species that is at risk of extinction, but the law that was designed to protect endangered species lacks the tools to deal with the threat."

More animals may also be called as witnesses in the fight against U.S. climate change policy, as the polar bear is far from the only animal threatened by the shrinking field of arctic ice. Ice-dwelling mammals such as the Pacific walrus and several species of seals "are in even worse shape," says G. Carleton Ray, an environmental scientist at the University of Virginia, Charlottesville. CBD has already filed petitions demanding that the ribbon seal and Pacific walrus also be listed as threatened.

–DAN CHARLES

Dan Charles is a freelance writer in Washington, D.C.

MICROBIOLOGY

Bacteria Are Picky About Their Homes on Human Skin

Julie Segre is touring the microbial landscape of our body's biggest organ, the skin. In anticipation of a $115 million, 5-year effort by the U.S. National Institutes of Health (NIH), she's traveling from head to toe, conducting a census of some of the trillions of bacteria that live within and upon human skin. Although their project is just getting off the ground, Segre, a geneticist at the National Human Genome Research Institute (NHGRI) in Bethesda, Maryland, and her colleagues have already uncovered a surprising diversity and distribution among skin bacteria. And a few oddities have emerged, too: Microbes known mostly from soils like healthy human skin, living in harmony with us; and the space between our toes is a bacterial desert compared to the nose and belly button.

Segre's work on what bacteria live where "is cool stuff," says Steven Salzberg, a bioinformaticist at the University of Maryland, College Park. "We need to increase our own and the public's awareness of the diversity and quantity of bacterial species on our own skin. The more people are aware, the more we can do to control infection."

Bacteria and other microbes that colonize our skin and other tissues outnumber the human body's cells 10 to 1, forming dynamic communities that influence our ability to develop, fight infection, and digest nutrients. "We're an amalgamation of the human and microbial genomes," says Segre. Recognizing this, NIH last year designated the Human Microbiome Project as one of its two Roadmap initiatives (*Science*, 2 June 2006, p. 1355). Researchers will sequence the genomes of about 600 bacteria identified as human inhabitants and get a handle on the 99% of bacteria that defy culturing but thrive in the skin, nose, gut, mouth, or vagina. "You have to understand what is the normal flora in the healthy skin to understand the impact of flora on disease," says Kevin Cooper, a dermatologist at Case Western Reserve University in Cleveland, Ohio.

As a first step, Segre, NHGRI postdoctoral fellow Elizabeth Grice, and their colleagues have studied five healthy volunteers, swabbing the insides of their right and left elbows. The site chosen isn't as unusual as it sounds; people with eczema often develop symptoms there. To survey the full thickness of skin, the researchers also used a scalpel to scrape off the top cells. And to reach even deeper, they took small "punches" of skin, a procedure akin to removing a mole.

From all the samples, Grice, Segre, and colleagues pulled out 5300 16*S* ribosomal RNA genes, which vary from microbe to microbe. After lumping together the most similar 16*S* genes, they came up with 113 kinds of bacteria and identified these dermal residents by matching the 16*S* genes to those of known bacteria. (Segre described the results at a recent meeting at Cold Spring Harbor Laboratory, and they are being published online 23 May in *Genome Research*.) "That's a lot of diversity, a lot of different organisms," says Martin Blaser, a microbiologist at New York University, who has done a similar survey of microbes living on the forearm, also finding a lot of diversity.

CREDIT: BILL LOSH/GETTY IMAGES

More than skin-deep. DNA surveys of the belly button, inner elbows, and elsewhere reveal diverse microbial communities.

Yet just 10 bacteria accounted for more than 90% of the sequences. Almost 60% of the 16*S* genes came from *Pseudomonas*, Gram-negative bacteria that flourish in soil, water, and decomposing organic debris. The next most common one, accounting for 20%, was another Gram-negative soil and water bug, *Janthinobacterium*. Neither had been considered skin microbes before this census. Although there were some differences among the volunteers in the microbes present, their elbows did share a common core set of microbes, the group reports.

The three sampling methods yielded slightly different results, with "punches" revealing a surprising number of bacteria under the skin—1 million bacteria per square centimeter compared with 10,000 from the scrapes. "I would have thought under the skin there would be fewer," says Salzberg.

Segre and her team have also begun sampling 20 other skin sites, including behind the ear and the armpit, from the bodies of volunteers. Skin varies in acidity, temperature, moisture, oil accumulation, and "different environments select for different microbes," says Blaser. Bacteriawise, reports Segre, "no subsite is identical."

Some researchers suspect that shifts in the makeup of skin microbial communities activate the immune system to cause diseases such as eczema. "If you know what the [healthy] flora is, then one strategy is to recolonize the area with the right flora," says Cooper.

–ELIZABETH PENNISI

A New Great Lake—or Dead Sea?

Turkmenistan intends to create a huge lake in the desert by filling a natural depression with drainage water. Critics say it's a bad idea that could even spark a war

ASHGABAT, TURKMENISTAN—Bone-dry and as forbidding as California's Death Valley, the windswept, 120-kilometer-long Karashor Depression—a natural bowl speckled with the ash-gray, mica-laden sand that gives the Karakum, or "Black Sand," Desert its name—might seem the last place in the world to put a lake. But on a fine day in October 2000, some 450 kilometers south of Karashor, President Saparmurat Niyazov leaned against a spade and breached a few-meters-wide earthen dam. Laborers took over, and soon water was gushing into the initial segment of a canal intended to fill Karashor to its rim. Golden Age Lake, the late president said, would become "the symbol of revival of the Turkmen land," covering 3500 square kilometers—nearly the area of Utah's Great Salt Lake.

With that gesture, Niyazov—known as Turkmenbashi, or "Father of the Turkmen People"—launched one of the most grandiose water projects ever undertaken. According to the plan, two canals that bisect the country will funnel runoff from heavily irrigated cotton fields into Karashor. The $6 billion project is designed to drain swamps and combat the buildup of salt and other minerals that have degraded three-quarters of Turkmenistan's arable land and eroded renowned archaeological monuments. "The lake will solve many problems," says Paltamed Esenov, director of the National Institute for Deserts, Flora, and Fauna in Ashgabat. Turkmen officials predict that the project will reclaim 450,000 hectares of waterlogged agricultural fields and create a habitat for migratory birds and an inland fishery.

Online
sciencemag.org
More on this story in *Science*'s Podcast

Next month, Turkmen engineers say they will complete the mammoth effort's first phase: excavation of the two "collector" canals, each hundreds of kilometers long. Water apparently has already begun trickling into Karashor. "We are carrying out a unique, pioneering project," says a senior engineer at the Turkmen State Water Research, Production, and Design Institute in Ashgabat, which leads construction of Golden Age Lake. "Everything we are doing is aimed at increasing agricultural productivity," says the engineer, who requested anonymity after agreeing to be interviewed without permission from Turkmenistan's Ministry of Foreign Affairs.

But Golden Age Lake has unleashed a torrent of criticism as well. "There's no sense in this," says Timur Berkeliev, a geochemist who coordinates the Worldwide Fund for Nature's Econet project in Turkmenistan. He and others are skeptical of plans to purify the runoff, laden with pesticides and fertilizers, and contend that the lake will become an artificial Dead Sea. "Trying to find value in this lake may be like trying to put lipstick on a pig," says Michael Glantz, director of the U.S. National Center for Atmospheric Research's Center for Capacity Building in Boulder, Colorado. "A bad idea, even for the best of intentions, is still a bad idea." Some experts believe that runoff will be insufficient to fill the lake, as the drainage water will evaporate or seep into the desert through unlined feeder canals.

That prospect raises fears that the lake could trigger a water war. Some observers worry that to prevent Golden Age Lake from running dry and to dilute tainted water, Turkmenistan might top it off with fresh water from the Amu Darya, a river on the border with Uzbekistan to the north. Uzbeks rely on the river for irrigation, and their leaders have said they would not tolerate a reduced share of the Amu Darya. "The lake project has incredible geopolitical implications," says Johan Gely, who works on water issues in central Asia for the Swiss Agency for Development and Cooperation. The senior water engineer insists such fears are unfounded: "Every drop of the Amu Darya is valuable, and nobody is planning to use this water for Golden Age Lake," he says.

Some see a window of opportunity to coax Turkmenistan to reconsider. Niyazov died in December 2006, and his successor,

CREDIT (INSET): JACQUES DESCLOITRES, MODIS LAND RAPID RESPONSE TEAM, NASA/GSFC

Making a lake. Two cross-country canals will funnel drainage water from Turkmenistan's heartland into the Karashor Depression.

Gurbanguly Berdimuhamedov, has not yet spoken publicly about the project. Foreign leaders have remained mum as well, perhaps in deference to Turkmenistan's growing clout as owner of the world's fifth largest natural gas reserves. In the meantime, Berdimuhamedov has promoted a gradual opening of the isolated country. "The leadership is now sensitive to world opinion," says Berkeliev. There might be one last chance, he says, to persuade authorities to convene an international scientific review before irreversible steps are taken to fill the lake. "This is the right time to do something," he says.

Back in the USSR

Centuries ago, central Asians learned how to make the most of the region's scarce water with networks of underground canals that conserved water for irrigation and drinking. "The tragic irony is that this region was home to one of the largest and most efficient irrigation systems in history, until the Mongol invasion destroyed much of the network," says Peter Sinnott, director of the Caspian Project at Columbia University.

Josef Stalin managed to outdo the Mongols. During the Cold War, when central Asia was part of the Soviet Union, Stalin's water managers cooked up a notorious fiasco. In the 1950s, they began to divert massive amounts of water from the Syr Darya into a network of canals to irrigate cotton fields in Uzbekistan. The Syr Darya is one of two main sources of water for the landlocked Aral Sea; the river's reduced flow resulted in the Aral's shrinkage to less than a quarter of its original surface area.

Soviet planners were pushing cotton in Turkmenistan as well, and in 1954, work commenced on the Karakum Canal, which would feed water from the Amu Darya—the other big Aral Sea source—into the Turkmen heartland. At 1375 kilometers in length, the Karakum waterway, completed in 1988, is the world's longest irrigation canal. It has been a boon for agriculture—it tripled the arable land in its vicinity—and provides water to the capital, Ashgabat.

But it has a dark side: A sizable fraction of the water that enters the canal (15% to 50%, depending on whom you ask) seeps through its unlined bed into the surrounding soil. The hemorrhaging created a patchwork of ponds and swamps and has exacerbated salinization. As the ground became waterlogged, the water table rose, bringing salts—primarily sodium sulfate—to the surface by capillary action. With evaporation, the brine crystallizes into mirabilite, a corrosive mineral that ruins oases and poisons fields. "Several kilometers to the left and right of the canal is a death zone," says a Turkmen government scientist who asked to remain anonymous to keep his job. "If you step in the extremely salty water, your shoes are destroyed within a week," adds a Western technician in Ashgabat who has visited the construction site of Golden Age Lake.

The Karakum Canal is not the only villain in the salinization saga. In the mid-1970s, Soviet engineers constructed drainage canals to discharge runoff into the desert. Dumping, coupled with overirrigation of farm fields, has saturated the ground and brought salt to the surface across the watershed. The water table is so high in the Dashoguz region, researchers say, that dozens of saline lakes have formed from water burbling up from the ground. "About 80% of arable land is damaged to different degrees," says Berkeliev. Many Turkmen farmers soak fallow fields in winter, wrongly believing that as fresh water seeps into the soil, it takes salt with it. "But this has the opposite effect," concentrating mirabilite, Berkeliev says: "This is a very complex problem, and the level of study is not adequate."

That hasn't stopped Turkmen authorities from forging ahead with a solution: the resurrection of a 1970s idea to divert Turkmenistan's irrigation runoff into Karashor, near the border with Uzbekistan. Niyazov dusted off a Soviet rough blueprint for an artificial lake, Glantz and others assert, as a strongman's way of showing dominion over nature. "Only a powerful state can build such a gigantic thing," Niyazov said in 2003. Turkmenistan's leader from the country's independence in 1991 until his death, Niyazov was anointed by parliament as Saparmurat Turkmenbashi the Great and, in 1999, made president for life. Golden Age Lake was not put to public consultation or debate. "It was almost impossible to object before," says Berkeliev. In 2004, after merely asking whether the project included ecological expertise, the country's sole homegrown environmental group, the Katena Ecological Club, was shut down.

One potential beneficiary of the lake project is the region's archaeological treasures. "Water and salt are the main enemies of archaeological sites," says the government scientist, who says that farmland and runoff have begun to encroach on what might be Turkmenistan's most famous site, the Bronze Age ruins of Gonur Depe (*Science*, 3 August 2007, p. 586). Salinization has already taken a heavy toll at one ancient monument, Little Kyz Kala in the medieval city of Merv, which has deteriorated especially rapidly in recent decades. The water table rose, soaking the foundations of the 1400-year-old brick fortress with salt and weakening them (see photos, below). With archaeologist Tim Williams and colleagues at University College London, Sébastien Moriset's team at the International Centre for Earth Construction of the Grenoble School of Architecture in France has helped Turkmen conservators improve drainage and apply sacrificial soil layers at monuments that will bear the brunt of erosion rather than the original walls.

Going, going ... The 1400-year-old Little Kyz Kala fortress in Merv was in bad shape in 1950 (*top*); a rising water table accelerated the erosion, greatly diminishing the monument by 2003.

CREDITS: (TOP TO BOTTOM) V. ITKINE; COURTESY OF TIM WILLIAMS/UCL; LOUISE COCKREUIL

Draining the runoff water from the landscape should, in theory, ameliorate salt-induced erosion of the monuments, says the government scientist. "How it will work in practice," he says, "we don't know."

Salvation or damnation?

To turn a dusty depression into a lake requires a whole lot of moisture. So the first and perhaps most formidable task was to excavate the two cross-country

Soaking up new ideas. Turkmen State University students are riveted by the words of a foreign lecturer.

The End of an Intellectual Dark Age?

ASHGABAT, TURKMENISTAN—This autumn, 80 top university graduates in this central Asian nation will take part in a revived system of candidate (the Russian equivalent of a Ph.D.) and doctoral degrees in fields as diverse as art history and zoology. If that sounds modest, consider how many students last year began postgraduate studies in Turkmenistan: zero. This is the country's first crop of postgrads since 1997.

That year, the nation's authoritarian former leader, Saparmurat Niyazov, abolished advanced degrees. Other elements of his stultifying program included halving undergraduate education to 2 years and lopping a year off secondary school. Niyazov also closed the Academy of Sciences in 1997, citing "the lack of any practical scientific results." Perhaps most insidious of all, his underlings enforced rote memorization of a book—the *Rukhnama*, a banal spiritual primer that Niyazov himself penned—as dogma.

Since Niyazov's death in December 2006, his successor, Gurbanguly Berdimuhamedov, has made education reform a top priority. He has upped university education to 5 years—six for aspiring physicians—and reinstated the lost year of secondary school. Science is back in fashion: "Science plays the leading role in the strong state, and therefore we should keep pace with

collector canals. Specialists plotted out routes that would make best use of natural topography. "In some places we had to dig as deep as 50 meters," says the senior water engineer. In other areas they built platforms or added boulders as obstacles to suppress the flow rate. When they encountered giant stone slabs, they invented equipment that could be inserted in cracks between layers to lift the rock out. Blasting was considered too expensive, and "we don't have reliable professionals for that purpose," says the senior water engineer.

The crew dug the northern canal in the Dashoguz region wider and deeper to allow for a larger water flow. For about half its length, the 432-kilometer Dashoguz Collector follows the bed of the ancient Uzboy River. The 720-kilometer Great Turkmen Collector starts in the Lebap region in the east and links up with the Dashoguz Collector 75 kilometers upstream of Karashor. About 45 kilometers from the depression, engineers built a 30-meter-tall, 600-meter-long dam to steer the water; otherwise it would have followed the lower-elevation Uzboy riverbed to the Caspian Sea. The senior water engineer says his engineers have also done some "sculpting" of Karashor's contours.

Water is now moving the length of the Dashoguz Collector and beginning to flow in the Great Turkmen Collector, the senior water engineer says. Satellite images confirm this. "It looks like canals, even unlined, can convey the drainage flow," says Leah Orlovsky, a water researcher at Ben-Gurion University of the Negev in Israel who works in Turkmenistan. On a flight from Tashkent to Tel Aviv last October, Orlovsky noticed that an area of roughly 20 to 25 square kilometers at the southern end of Karashor was flooded.

Filling the lake should take several decades, says Esenov of the desert research institute. Water must first flow into the capillaries—a 1000-kilometer network of small feeder canals linking at one end to agricultural drainage ditches and at the other to small reservoirs or to the vast collector canals. Pumping stations regulate the flow into the collectors. Eventually, the senior water engineer says, the groundwater table should drop by a couple of meters, allowing for the gradual desalinization and reclamation of farm fields.

Although Golden Age Lake could save some iconic monuments, lesser known archaeological sites were damaged during construction. "They just bulldozed some small monuments and sites that hadn't been excavated yet," says the government scientist. The project's design called for an archaeology rescue program, he says, but it had no funds. Living heritage is being lost as well. The collectors have raised the water table along their length, spoiling drinking water wells in some desert settlements. "Villages with ancient roots are being moved," says the independent scientist. "It's a degradation of the cultural landscape."

Future plans call for widening and deepening both collectors, says the senior water engineer. But there are no plans to line them. He referred questions about their dimensions and anticipated flow rates to institute colleagues, who were not available for interviews. One told *Science* privately that the lake's depth should reach 130 meters and its anticipated volume is 135 to 145 cubic kilometers.

"Data from Turkmenistan are hard to come by … and not so reliable," says Glantz. But even rough approximations suggest that the project is doomed, says Berkeliev. The quality of the lake will depend on what goes into it, and Turkmen authorities in the past have predicted a water inflow of 10.5 cubic kilometers a year. About two-thirds will come from Dashoguz, including cross-border runoff from the Khorezm region of Uzbekistan; the Great Turkmen Collector will supply the other third of the water. However, Uzbekistan plans to build a drainage canal from Khorezm to the Aral Sea, so the amount feeding Golden Age Lake would eventually taper off, says Kai Wegerich, a central Asia water expert at Wageningen University in the Netherlands. "If the Uzbek drainage canal is built, it might not make sense anymore to construct the lake," he says.

Khorezm canal or no, Berkeliev says his calculations are damning. Based on the high evaporation rate in Karakum, he asserts, "there will never be a water body there." Others say Golden Age Lake may well come into being but is fated to become an environmental nightmare: a salty broth of organic pesticides and fertilizers.

Not so, says Vyacheslav Zharkov. He and his colleagues at the desert research institute in Ashgabat are devising filter media that absorb heavy metals and organic contaminants from runoff. These can be installed at treatment plants at points where water enters the collector canals—if the Turkmen government finds money to build such treatment plants. "After treating water with our sorbents, it is suitable for agriculture and for drinking," Zharkov says. He claims that salt will be drawn out as water moves along the

CREDIT: MITSUM STONE

its latest achievements," Berdimuhamedov said recently, according to the state press. But recovery will not be easy. "After so many years of the forced degradation of the education system, it's really hard to revive it," says one Turkmen government scientist. "The serious scientists didn't wait for changes within the country—they left."

Turkmenistan is not the first modern nation to willfully erode its intellectual capacity: Afghanistan under the Taliban, for instance, suffered severely. But Turkmenistan's descent took place largely out of sight, as Niyazov isolated the country and placed sharp limits on international cooperation. In this twilight, in 2001, the *Rukhnama* appeared. The book is a mix of folksy guidance about how to lead a good life and a history of the Turkmen people that mangles the chronology of real events and fabricates others. "It did great damage for historians," says the government scientist. Workplaces formed *Rukhnama* study circles, and TV programs showed children reciting passages while professing their love for Niyazov. *Rukhnama* knowledge was necessary to pass exams, including the driver's test.

The *Rukhnama* is still for sale in Ashgabat, and in some primary and secondary schools "it remains a strong part of the curriculum," says Leon Yacher, a geographer at Southern Connecticut State University in New Haven, who lectured in Turkmenistan last month. When he visited a school in Turkmenbashi, a city on the Caspian coast, "every student had a copy of the book on their desk, and they were expected to read from it every day." But to the relief of scholars, the *Rukhnama* is being phased out in universities and government offices.

Turkmen academics are trying to pick up the pieces. "A change of the curriculum is needed badly," Yacher says. One problem is that there are few solid Turkmen textbooks, and no recent textbooks in Russian or in English, says the government scientist. That matches the general decrepitude of the faculty. Even in a field that was in favor under Niyazov—archaeology—the department was eliminated at Turkmen State University in 1999 and, says the government scientist, "the youngest archaeologist we have is a 60-year-old guy. When the last generation of archaeologists is gone, only foreigners will work here."

Among signs of progress, construction has begun on a $35 million building for Turkmen State's physics and mathematics faculty, and a new campus is in the works for Turkmen State Medical Institute. The country is looking beyond its borders as well, with plans this fall to dispatch 1500 students to overseas universities, including Columbia University. "If [students] are off-the-charts good, we should do what we can to overcome any obstacles and get them here," says Peter Lu, a physicist at Harvard University, who lectured in Turkmenistan in 2005. Foreign institutions can play a critical role in the intellectual revival, starting with the next generation. **–R.S.**

canals. "We have asked how the salt will be removed. They say the water will clean itself. Nobody is able to explain to me how this works," says the Western technician. Berkeliev too says he is mystified.

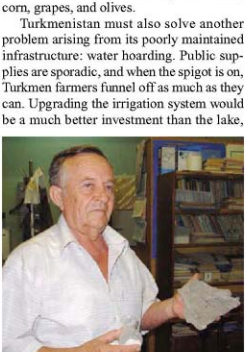
Looming shortage

The overarching question is whether Turkmenistan might tap the Amu Darya to improve the new lake. Under the Soviet-era water-sharing agreement, Turkmenistan and Uzbekistan each can use up to 22 cubic kilometers of water flowing out of Afghanistan and along their shared border—despite a huge difference in population size. (Turkmenistan has 5 million people; Uzbekistan has 28 million.) "The Uzbeks will not tolerate any 'vanity diversions' to the new lake," says Glantz. In a tense situation, "new diversions will lead to a real war."

Even if water isn't diverted to the lake, Afghanistan's plans to rev up irrigation are likely to curtail the Amu Darya's flow. Currently, it uses only a few cubic kilometers each year. "They are planning a massive expansion of irrigation," says Wegerich. Several major projects launched in the last 2 years aim to irrigate more than 1 million hectares, with completion dates staggered over the next 5 to 15 years. Adds Glantz, "The Uzbeks think it is decades away. Wrong." A complicating factor is the retreat of glaciers in the Pamir Mountains—the source of much of central Asia's fresh water. "Eventually, there will be no Amu Darya, no Syr Darya," Mamadsho Ilolov, president of Tajikistan's Academy of Sciences, told *Science*. Golden Age Lake, he says, "will be very dangerous for neighboring countries."

The best solution to Turkmenistan's water problems, Berkeliev and others argue, is conservation. Currently, Turkmenistan uses 5000 cubic meters of water per capita per year. That's twice the rate of Uzbekistan and more than 10 times that of Israel. "We are the champions of water waste," says Berkeliev. It's high time, he and others say, that the country revises its Soviet-era agricultural system and switches to water-saving technologies, like drip or subsoil irrigation, and converts a significant portion of farmland to less water-intensive crops like wheat, corn, grapes, and olives.

Turkmenistan must also solve another problem arising from its poorly maintained infrastructure: water hoarding. Public supplies are sporadic, and when the spigot is on, Turkmen farmers funnel off as much as they can. Upgrading the irrigation system would be a much better investment than the lake, says Aral Sea expert Philip Micklin, a geographer at Western Michigan University in Kalamazoo. In his view, Golden Age Lake is "a big waste of money." When the plan was being put together in the late 1990s, it had a conservation component—"but that disappeared," says the independent scientist. "If they spent half the budget of the lake on water conservation," he says, "they would not have had to build the lake."



CREDIT: MATTHIAS STONE

Soaking up contaminants. Vyacheslav Zharkov says his sorbents can render the waters of Golden Age Lake suitable for drinking.

The senior water engineer says he is not bothered by the criticism and that it will not derail the lake project. "We faced the same opposition when we built the Karakum Canal," he says. "Any such great project will have negative effects. But these are outweighed by the benefits."

Berkeliev says it's refreshing to be able to have this debate; it could never have happened under Niyazov. "But to change the minds of decision-makers, we need strong support from the outside," he says. "We must have an international review of this project while there's still time," adds geographer Igor Zonn of the Engineering Research Center on Water Management, Land Reclamation, and Environment in Moscow. That might be possible, as Turkmenistan continues a cautious opening up to the world. "We are trying to increase international cooperation on environmental issues," Ogulsona Karyeva of the Ministry of Nature Protection told a Fulbright conference in Ashgabat last month.

"We would be very happy to work with foreign scientists," says Esenov. "It's a complex problem." That's something everyone can agree on.

–RICHARD STONE

Farm labor. Leaf-cutter ants tend their fungus garden, a complex miniature ecosystem.

COMMUNITY ECOLOGY

All That Makes Fungus Gardens Grow

The discovery of a parasitic yeast draws attention to the ways that pathogens can stabilize ant agriculture and other symbiotic networks

Fifty million years ago, while the earliest primates were still scurrying from tree to tree, scrounging fruits and insects, attine ants were growing their own food. They were so adept at domesticating mushrooms that hundreds of species have descended from the original farmers, all of them cultivating fungi.

Humans could learn a lot from the ants' success. Over the past 10 years, researchers have come to realize that the fungus gardens thrive because of an intricate web of bacteria and fungi that includes both pests, such as a newly discovered black yeast, and partners, including bacteria that keep pathogens in check. By studying these relationships, biologists hope they'll uncover lessons about the evolution of such interactions, knowledge that will help humans better manage microbes in medicine and agriculture. "It's a system that works," says John Morrissey, a microbiologist at University College Cork in Ireland. "If you could develop a bacterial inoculant that was as successful in controlling a specific pathogen as the [beneficial bacteria] are for the ants, you'd be on to a real winner."

Complex network

Ant agriculture runs the gamut. For leaf-cutter ants, farming is big business. They're the most notorious of the more than 230 described species of fungus gardeners, forming colonies of millions of workers that can defoliate a tree or crop in mere hours. The ants use the harvest to fertilize hundreds of separate fungus gardens in an elaborate subterranean compound. Most attine gardens, however, are small-scale operations: Their inconspicuous colonies are tended by as few as a dozen workers that scavenge bits of detritus to feed a spongy handful of fungus.

But from the most primitive gardener to the dreaded leaf-cutter, all attines would starve if deprived of their fungal crops. When an ant queen leaves home to mate and found a new colony, she must take a little mouthful of the fungus with her to start a garden.

Although naturalists have known since 1874 that the attine ants are fungus gardeners, more than a century passed before Ph.D. student Cameron Currie began to chip away the microbial complexity underlying the ant-fungus symbiosis. While at the University of Toronto in Canada, he discovered that ant gardens often contained a second fungus, *Escovopsis*. When he grew it on culture plates with different food sources, Currie determined that *Escovopsis* is a pathogen with a sweet tooth for only the ants' cultivar. What's more, the pathogen's evolutionary tree had the same basic shape as those of the ants and their crop, indicating that all three had coevolved since the beginning of ant agriculture, Currie and colleagues reported in 2003 (*Science*, 17 January 2003, pp. 325, 386).

Although Currie isolated *Escovopsis* from up to 75% of the gardens of several attine species in Panama, this pathogen rarely seemed to do much damage. The reason, it turned out, was a fourth symbiont: Currie found that actinomycete bacteria, housed and nourished in pits on the ants' bodies, produce chemicals that keep *Escovopsis* in check.

The four-part garden symbiosis of ant, cultivar, pathogen, and bacteria interrupted a scientific tradition of studying symbionts two at a time—think corals and algae, for example, or soybeans and nitrogen-fixing bacteria. The discovery accelerated a transition toward thinking of interacting organisms in trios or networks, not pairs.

"When I got into this stuff, it was two symbionts," says Ted Schultz, an entomologist at the Smithsonian National Museum of Natural History in Washington, D.C., who has studied attine evolution for nearly 30 years. "I was stunned" when Currie identified two more.

Now, in a paper in this month's issue of *Ecology*, Currie and his colleagues introduce a fifth symbiont. "[It] just continues the trend of being repeatedly surprised by how complex this system is," Schultz says.

The first hint of the new player came when Currie, now at the University of Wisconsin (UW), Madison, cultured the actinomycete bacteria from an ant called *Apterostigma*. In addition to white bacterial spots, a black yeast often appeared on the same culture plates. Ainslie Little, now a postdoctoral fellow at UW Madison, took a closer look at the yeast. She treated worker ants in a vial coated with a selective antibiotic that would rub off on the ants and kill the yeast but not the other symbionts.

At first, it looked like the experiment was a bust: Getting rid of the yeast had no effect on the ants or their crop. But when Little spritzed half the ants' gardens with a solution of *Escovopsis* spores, the yeast suddenly revealed its true colors. Over 3 days, ants with black yeast infections lost twice as much of their crop to *Escovopsis* as the yeast-free ants. When Little grew the actinomycetes in petri dishes with the yeast, the yeast ate the bacteria, demonstrating that they rob ants of an important defense against *Escovopsis*.

DNA studies showed that the black yeasts are widespread on the attine ant family tree, thriving near the pits where the ants

CREDIT: ALEX WILD 2006

house the actinomycetes. Like the cultivar, *Escovopsis*, and actinomycetes, the yeast has been part of the attines' microbial balancing act since the ants first began to farm, Currie says.

"It's really exciting," Schultz says of the fifth symbiont. And Ulrich Mueller, an integrative biologist at the University of Texas, Austin, agrees: It's "interesting to what extent the presence of [another] symbiont can fundamentally change the interaction of two other symbionts."

No cheating allowed

Currie thinks that understanding three-, four-, and five-way interactions like the ones in the fungus gardens may ultimately revolutionize the way we think about the evolution of mutualism. Why two parties should cooperate—whether it's two species over evolutionary time or two people over the course of a day—is one of science's big mysteries (*Science*, 1 July 2005, p. 93). In a two-player partnership, cheaters should be able to get ahead by reaping benefits without paying their dues, destabilizing the agreement. But the ants have lived stably with two mutualists for millennia.

Moreover, the two antagonists rarely get too far out of line. Diseases and pests of humans and their crops, on the other hand, have evaded control measures in a matter of decades. So everyone, from crop scientists to basic evolutionary biologists, is itching to know the secrets to both cooperation and control in the ants' gardens.

Morrissey, for example, would like to incorporate beneficial microbes into human agriculture to reduce chemical input. The attine system "shows it is possible to set up a structured [microbial] community … over a very long term" for biological control, he says.

The "key question" for figuring out how to manage beneficial microbes, says R. Ford Denison, a crop ecologist at the University of Minnesota, St. Paul, is why the ants' actinomycetes never turn against their hosts. Antifungal compounds are expensive to make, and selfish bacteria that don't make the compounds should be able to reproduce faster and eventually overrun the helpers. But somehow the ant system is robust to cheaters.

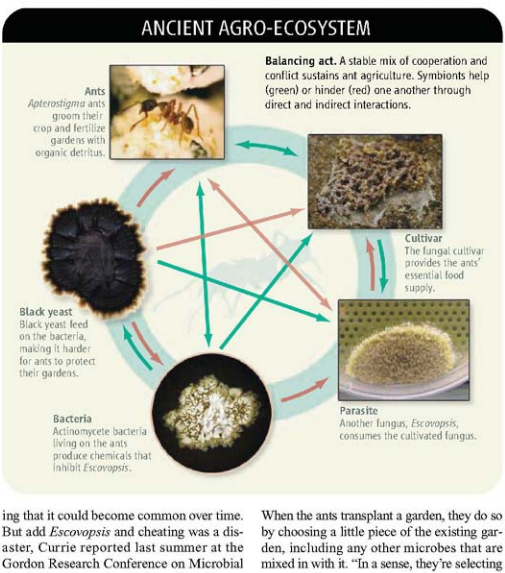
This is where the multipart interactions come in, Currie says. He thinks there is a reason that each partnership in the garden has a parasite: *Escovopsis* intrudes on the ant-fungus mutualism, and the yeast disrupts the ant-actinomycete mutualism. These parasites may be the very thing that keeps the mutualists cooperating, Currie adds.

Currie and his colleagues tested this idea by forcing the fungus or the ants to cheat on each other. A selfish fungus would reproduce more and feed the ants less. To simulate this, Little removed specialized nutrient-laden structures from much of the cultivar. To shortchange the fungus, she reduced the proportion of fungus-tending workers in the colony.

With no *Escovopsis* around, the negative effects of cheating were minimal, suggesting that it could become common over time. But add *Escovopsis* and cheating was a disaster, Currie reported last summer at the Gordon Research Conference on Microbial Population Biology in Andover, New Hampshire. When either the ants or the fungus were cheating, *Escovopsis* took over more of the garden, killing the fungus and leaving the ants with little to eat. Parasites, Currie says, may play a crucial and underappreciated role in keeping cooperators honest by raising the costs of cheating.

But that doesn't explain what keeps the parasite itself from overrunning the fungus garden, the same way human pathogens have outpaced antibiotics. New results from Currie's lab, not yet published, indicate that *Escovopsis* does evolve resistance but never gets ahead because the actinomycetes themselves are so adept at evolving new antifungal compounds. And if they don't do it fast enough, the ants can acquire new actinomycete strains.

But other organisms, yet to be described, may also play a part in stabilizing fungus garden ecology. "There are additional biofilms associated with [the fungus], stuff that grows on the fungus, or in the substrate, wherever anything else can move in," says Mueller.



When the ants transplant a garden, they do so by choosing a little piece of the existing garden, including any other microbes that are mixed in with it. "In a sense, they're selecting on an entire community that has desirable properties," Mueller says. The idea still needs to be tested. Meanwhile, only the daring would place bets on how many symbionts have yet to turn up. Currie expects a few more; Mueller predicts hundreds. Evolutionary biologist Jacobus Boomsma of the University of Copenhagen in Denmark says that regardless of the number of symbionts, the fungus gardeners are poised to answer critical questions about cooperation, conflict, and microbial ecology. "This system," he says, "will keep inspiring us for at least a decade more."

–ELSA YOUNGSTEADT

Final touches. Technicians at Cape Canaveral ready GLAST for attachment to its launch vehicle.

ASTROPHYSICS

GLAST Mission Prepares to Explore The Extremes of Cosmic Violence

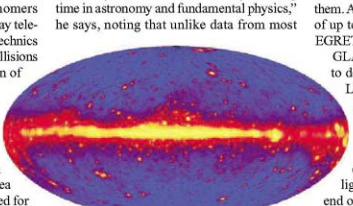
NASA's new gamma ray observatory will probe the most energetic radiation ever studied, the product of cataclysmic events deep in space

In July 1967, U.S. surveillance satellites looking for signs of a Russian nuclear test in space recorded two flashes of gamma radiation. Scientists quickly determined that the high-energy bursts did not come from a nuclear explosion, which would have generated a more sustained stream of gamma rays and also produced lower energy radiation detectable by other satellite instruments. Only years later did they realize that the flashes—named gamma ray bursts (GRBs)—originated in violent events deep in space. In scanning the heavens for an enemy secret, they had stumbled upon a cosmic one.

That serendipitous discovery opened a window on previously unknown phenomena whose signatures lay at the gamma end of the energy spectrum. Since then, astronomers have dispatched a number of gamma ray telescopes into space to glimpse the pyrotechnics unleashed by violent events such as collisions between neutron stars and the emission of particle jets by massive black holes.

Now, researchers are opening the window wider with a new telescope designed to record gamma radiation several orders of magnitude higher in energy than current instruments can detect. NASA's Gamma-ray Large Area Satellite Telescope (GLAST), scheduled for launch next month, will also be the first instrument of its kind to survey the entire sky several times a day, increasing the chances of finding and following extreme astronomical phenomena anywhere in the universe.

Researchers say GLAST's powerful combination of sensitivity and sweep will yield a rich harvest of data that could answer a host of astronomical questions such as how supermassive black holes behave and how cosmic rays originate. A tantalizing possibility is that observations from GLAST will help physicists discover the fundamental nature of dark matter, which makes up 10 times as much of the universe as the familiar matter of planets and stars. The mission represents a convergence of the quests to understand the large and the small, says Rene Ong, an astronomer at the University of California, Los Angeles. "We are just at the start of a very exciting time in astronomy and fundamental physics," he says, noting that unlike data from most previous missions, GLAST's will be available in real time to scientists—and the public—anywhere in the world starting a year after the launch.



Coming attraction. Simulated "gamma ray sky" shows how new observatory will view the universe.

CREDITS: TOP TO BOTTOM: JIM GROSSMANN/NASA; NASA

Catching rays

Built over a decade at a cost of $690 million, GLAST is a feat of engineering. Its main instrument is a 3-ton detection system called the Large Area Telescope (LAT), which consists of two devices to track the direction and energy of incident gamma rays—energetic photons of extremely high-frequency electromagnetic radiation. The direction detector is a four-by-four matrix of towers that are essentially layers of tungsten and silicon stacked one on top of another. When a gamma ray slams into a tungsten layer, there's a chance its energy will be transformed into an electron and a positron that are propelled along the path the ray would have taken. As the two particles travel through silicon layers in the stack, they generate currents that reveal their direction. When they emerge from the bottom of the stack, the particles enter a chamber of cesium iodide—the telescope's energy-detecting device—producing a flash of light whose intensity shows how fast they had been moving and thus the energy of the gamma ray.

The stacked design gives LAT a much larger collecting area than previous gamma ray telescopes such as the Energetic Gamma Ray Experiment Telescope (EGRET), which flew on NASA's Compton Gamma Ray Observatory from 1991 to 2000. More collecting area means increased chances of a collision. And that's exactly what's needed in order to detect higher energy gamma rays, explains Steven Ritz, GLAST project scientist at NASA's Goddard Space Flight Center in Greenbelt, Maryland, because they are so rare that a less sensitive instrument would miss them. As a result, LAT can detect gamma rays of up to 300 billion electron volts, 10 times EGRET's upper limit (see figure, p. 1009).

GLAST has a second instrument designed to detect lower energy gamma rays that LAT would not register. Called the GLAST Burst Monitor (GBM), it's a set of 12 sodium iodide disks and two bismuth germanate disks pointed in different directions, covering practically the entire sky. The disks produce light when struck by photons at the lower end of the gamma spectrum; scientists can trace the direction of the incident rays simply by noting which disk bears the brunt of the collision. GBM will detect rays between 10

KeV and 25 MeV, overlapping with LAT's lower limit of 20 MeV. "Together, the two instruments give us vast energy coverage," says Ritz, adding that "if GLAST were a piano, it would have 23 octaves."

One of the challenges in designing the system was to ensure that it would run on the small amount of power available from the satellite's solar panels. Robert Johnson, a physicist at the University of California, Santa Cruz, who led the engineering of the tower array, says simplifying the electronics was part of the solution. "We ended up at 160 watts," he says. "That's a couple of light bulbs of power for over 900,000 channels."

Black hole, bright lights

Astronomers will be eagerly scanning GLAST data for clues to what goes on near monstrous black holes that sit at the centers of galaxies. Such objects can be as massive as hundreds of thousands or even billions of stars. As their enormous gravity sucks matter into a whirling disk around them, opposing jets of particles shoot away from their poles at nearly the speed of light. If a jet from such an active galactic nucleus (AGN) happens to be pointed at Earth, astronomers call it a blazar. The process generates radiation across the entire electromagnetic spectrum, including high-energy gamma rays.

The Compton Observatory identified 66 blazars during its time in orbit. Astronomers have since puzzled over how these beasts accelerate particles to such high speeds. GLAST is expected to find thousands of new blazars because of its sensitivity and periodic surveying of the sky, and the data it sends back should provide a sharper, more dynamic picture of these events than Compton did, says Ritz. He expects blazars and AGN to be a "bread and butter" topic for researchers analyzing GLAST data.

Alan Marscher, an astronomer at Boston University, agrees. Last month in *Nature*, Marscher and colleagues presented x-ray, radio, and visible light observations from a blazar, suggesting that the "accretion disk" spinning around the black hole had caused the magnetic field in the galactic center to coil into a spiral, leading to the emission of particulate jets from its core. He says GLAST will help him test the theory by observing how the brightness of gamma radiation from different blazars changes over time. "We expect time delays between the peaks in the flares at different gamma ray energies and relative to the flares in x-ray, visible-light, and radio emission," Marscher says. That signature, he says, could offer a deeper look into the heart of a blazar.

CREDIT: NASA (AD); GENERAL DYNAMICS

Outracing LHC?

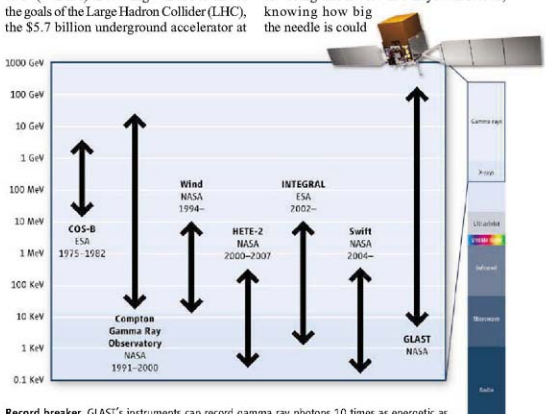
Researchers involved with GLAST call the mission a unique marriage between particle physics and astronomy. Some are hoping to justify that description in grand fashion in the years to come by carving out a prominent role for GLAST in finding the elusive particle that constitutes dark matter.

As its name implies, astronomers cannot see dark matter; only its gravity gives it away. Many theorists think it consists of still-unknown "weakly interacting massive particles" (WIMPs). Detecting WIMPs is one of the goals of the Large Hadron Collider (LHC), the $5.7 billion underground accelerator at CERN that is expected to come on line this summer. But if the hypothesized particles do turn up there, physicists will still need to confirm that they make up the dark matter out in space. That's where GLAST would come in.

According to theory, in the rare event when two WIMPs collide, they annihilate each other and give off gamma rays. Such collisions are most likely in galactic regions where dark matter is densely concentrated.

"We would look in those known directions to see if GLAST is picking up an excess of gamma rays," says Johnson, who started his career as a particle physicist before being completely "consumed" by the GLAST project a decade ago. He's now the co-convener of the mission's dark matter science working group. "The dream scenario is that we see the signature of dark matter before LHC turns on," he says with a chuckle. "But we'd be perfectly happy if they saw it first and determined the particle's mass, and then we went out and found it in space."

By the same token, observations from GLAST could help LHC in its quest to identify WIMPs, says Dan Hooper, a theoretical physicist at Fermi National Accelerator Laboratory in Batavia, Illinois. He says LHC will generate such a huge volume of data that any hint from GLAST about the energy released by an annihilating WIMP pair would help LHC scientists to focus their search. "If you are doing this needle-in-a-haystack search, knowing how big the needle is could be key," he says. However, Hooper cautions that gamma rays produced by WIMP collisions could turn out to be too faint for GLAST to see.



Record breaker. GLAST's instruments can record gamma ray photons 10 times as energetic as the detectors aboard earlier satellites could handle.

When the data start streaming in, researchers will be able to sink their teeth into them for insights into other fundamental problems. For example, says Neil Gehrels, deputy project scientist, GLAST might catch small, primordial black holes in a vanishing act, confirming a prediction by Stephen Hawking that such objects shrink by emitting radiation and eventually evaporate into a crematory flash of gamma rays.

"It could well be that the most interesting observations turn out to be something entirely new and unexpected," says Ritz. If that were to happen, as witnesses to the discovery of GRBs can testify, it would not be the first time.

–YUDHIJIT BHATTACHARJEE

LETTERS

edited by Jennifer Sills

Free Access to Landsat Imagery

Free image. This Landsat 5 image of the southeastern corner of the Black Sea is part of the general U.S. archive that will be accessible for free under the new USGS policy.

WE ARE ENTERING A NEW ERA IN THE LANDSAT Program, the oldest and most venerable of our Earth-observing satellite programs. With little fanfare, the U.S. Geological Survey (USGS) has begun providing imagery for free over the Internet. Throughout the history of the Landsat Program, the cost and access to imagery has always limited our ability to study our planet and the way it is changing. Beginning with a pilot program to provide "Web-enabled" access to Landsat 7 images of the United States that were collected between 2003 and this year, the USGS now plans to provide top-quality image products for free upon request for the entire U.S. archive, including over 2 million images back to Landsat 1 (1972) [for details and schedules, see (*1*)]. The release by NASA and the USGS in January 2008 of a new Landsat Data Distribution Policy (*2*) was a key step to this goal. Free imagery will enable reconstruction of the history of Earth's surface back to 1972, chronicling both anthropogenic and natural changes during a time when our population doubled and the impacts of climate change became noticeable.

THE LANDSAT SCIENCE TEAM: CURTIS E. WOODCOCK,[1]* RICHARD ALLEN,[2] MARTHA ANDERSON,[3] ALAN BELWARD,[4] ROBERT BINDSCHADLER,[5] WARREN COHEN,[6] FENG GAO,[5] SAMUEL N. GOWARD,[7] DENNIS HELDER,[8] EILEEN HELMER,[9] RAMA NEMANI,[10] LAZAROS OREOPOULOS,[5] JOHN SCHOTT,[11] PRASAD S. THENKABAIL,[12] ERIC F. VERMOTE,[13] JAMES VOGELMANN,[14] MICHAEL A. WULDER,[15] RANDOLPH WYNNE[16]

[1]Department of Geography and Environment, Boston University, Boston, MA 02215, USA. [2]University of Idaho Research and Extension Center, Kimberly, ID 83341, USA. [3]USDA–Agricultural Research Service, Hydrology and Remote Sensing Lab, Beltsville, MD 20705, USA. [4]European Commission Joint Research Center, Institute for Environment and Sustainability, Global Environment Monitoring Unit, 21020, Ispra, Varese, Italy. [5]NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA. [6]USDA Forest Service, Corvallis, OR 97331, USA. [7]Department of Geography, University of Maryland, College Park, MD 20742, USA. [8]Electrical Engineering and Computer Science Department, South Dakota State University, Brookings, SD 57007, USA. [9]International Institute of Tropical Forestry, U.S. Forest Service/Rocky Mountain Research Station, Fort Collins, CO 80526, USA. [10]NASA Ames Research Center, Ecosystem Science and Technology Branch, Moffett Field, CA 94035–1000, USA. [11]Rochester Institute of Technology, Chester F. Carlson Center for Imaging Science, Rochester, NY 14623, USA. [12]International Water Management Institute (IWMI), 127, Sunil Mawatha, Battaramulla, Colombo, Sri Lanka. [13]Department of Geography, University of Maryland, College Park, MD 20742, USA. [14]SAIC/USGS EROS, Sioux Falls, SD 57198, USA. [15]Canadian Forest Service, Pacific Forestry Centre, Victoria, BC V8Z 1M5, Canada. [16]Department of Forestry, Virginia Tech University, Blacksburg, VA 24061, USA.

*To whom correspondence should be addressed. E-mail: curtis@bu.edu

References

1. USGS Technical Announcement (http://landsat.usgs.gov/images/squares/USGS_Landsat_Imagery_Release.pdf).
2. Landsat Missions (http://ldcm.usgs.gov/pdf/Landsat_Data_Policy.pdf).

Why Rowe-Clark Doesn't Teach by the Book

WE ARE WRITING ON BEHALF OF THE 150 students attending Rowe-Clark Math and Science Academy, The Exelon Campus of Noble Street Charter School, and all of their teachers. *Science* reporter J. Mervis visited our campus for several hours this fall to observe our science and math classes, learn about our school, and interview students and teachers in preparation for the News Focus story "A new bottom line for school science" (22 February, p. 1030).

Rowe-Clark Math and Science Academy is an exemplary school, showcasing innovative, engaging science and math teaching and learning, as well as a model partnership between a school and a business. We thank Mervis for his attention to our school and the work we are doing with Exelon.

However, we feel that Mervis did a disservice to *Science* readers, as well as to Rowe-Clark students, teachers, and parents, due to a substantial—and even offensive—error in his reporting.

Mervis reported that freshmen taking physics at Rowe-Clark do not use a textbook. This is accurate. What is blatantly inaccurate is the reason Mervis cites: "because so many of them wouldn't be able to read it." V. Galarza explained clearly to Mervis her approach to teaching physics, and Mervis observed this during a 2-hour visit to our physics classroom. Using the modeling technique, Galarza guides students through hands-on experiments in which they collect data and, from that data, derive the laws of physics. Students remember what they learn because they have discovered it for themselves, not read it in a textbook. Many of our students are admittedly behind others at their grade level, but they are able readers who use textbooks in classes where that is deemed by the teacher to be the best tool for learning. In physics, active experiments make for more powerful and lasting learning than textbooks, and the modeling technique is an innovation

in science education that is becoming increasingly widespread. It is a student-centered instructional strategy through which students participate in active scientific inquiry, discourse, and evaluation of evidence.

RACHEL KRAMER AND VANESSA GALARZA

Rowe-Clark Math and Science Academy, 3645 West Chicago Avenue, Chicago, IL 60651, USA.

Science Education: Should Facts Come First?

IN THEIR EDUCATION FORUM "APPLICATION of Bloom's taxonomy debunks the 'MCAT myth'" (25 January, p. 414), A. Y. Zheng *et al.* suggest using Bloom's taxonomy as a tool for assessment of lower- and higher-level thinking. We think that Bloom's taxonomy should be considered more carefully before it is applied to the assessment and reform of undergraduate courses.

Bloom's taxonomy demonstrates a progressive sequence in human cognition, from simple (lower) to complex (higher) thought processes. However, it does not account for one important factor: a temporal or chronological sequence. At different stages of the whole educational process, instructional purposes are different. Assessments should reflect such purposes. Time-consuming lower-level fact instruction at earlier stages will contribute to and guarantee higher-level thinking. Without assessment of lower-level thinking (that is, the students' knowledge base and comprehension) at students' earlier stages, instructors may not know whether students are well equipped to advance to concepts that require higher-level thinking and how far the students could go.

All of the sources used by Zheng *et al.* were admission tests or first-year tests. The goal of this kind of test is to identify gaps in a student's mastery of basic facts, and higher-level questions are less effective in meeting this goal. It may be unrealistic and dangerous in assessment to jump to a greater proportion of higher-level thinking at the cost of possible ignorance of students' mastery of basic facts.

SHESEN GUO

Qianjiang College, Hangzhou Normal University, Hangzhou, Zhejiang 310012, China. E-mail: guoshesen@126.com

Response

GUO IS CORRECT IN POINTING OUT THAT THE explicitly hierarchical level of Bloom's taxonomy implies a chronological sequence in instruction (*1*, *2*). The literature does not, however, support the claim that factual recall should be the primary focus of early courses in a curriculum and that higher-order thinking should be reserved for later courses. Most educators believe that students should be working at higher levels in Bloom's hierarchy as often as possible. An author on both the original and revised Bloom's taxonomy (*1*, *2*) writes that educational objectives above the Factual Knowledge level "are usually considered the most important goals in education" [page 213 in (*3*)], including primary and secondary education. The recent revision of Bloom's taxonomy introduces a two-dimensional framework designed to assess how well the educational objectives from any course fulfill all elements in Bloom's framework (*2*, *3*).

In addition, we are not aware of data supporting Guo's assertion that "[t]ime-consuming lower-level fact instruction at earlier stages will … guarantee higher-level thinking." In our experience, an overemphasis on lower-order thinking at early stages of instruction impedes progress in later courses that also demand higher-order thinking. Problems arise because students have been trained to associate memorization with academic success.

Guo's letter reflects a view held by a substantial proportion of instructors that introductory science courses should focus primarily or even exclusively on factual content. The data in our Education Forum (25 January, p. 414) indicated that biology students who intend to pursue medical or graduate school are poorly served by such courses because the exams required for admission emphasize questions above the bottom rung on Bloom's taxonomy.

SCOTT FREEMAN, JANESSA LAWHORN, ALEX ZHENG

Department of Biology, University of Washington, Seattle, WA 98195, USA.

References

1. B. S. Bloom, Ed., *Taxonomy of Educational Objectives: The Classification of Educational Goals, Handbook I: Cognitive Domain* (David McKay, New York, 1956).
2. L. W. Anderson, D. R. Krathwohl, Eds., *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives* (Longman, New York, 2001).
3. D. R. Krathwohl, *Theory Pract.* **41**, 212 (2002).

Letters to the Editor

Letters (~300 words) discuss material published in *Science* in the previous 3 months or issues of general interest. They can be submitted through the Web (www.submit2science.org) or by regular mail (1200 New York Ave., NW, Washington, DC 20005, USA). Letters are not acknowledged upon receipt, nor are authors generally consulted before publication. Whether published in full or in part, letters are subject to editing for clarity and space.

A Victory for PETA

AFTER 2 YEARS OF LOBBYING BY PEOPLE FOR the Ethical Treatment of Animals (PETA), CareerBuilder has stopped its chimpanzee advertisement campaign. On 23 January 2007, PETA announced online that CareerBuilder had agreed to stop this series of advertisements; however, CareerBuilder did not sign a pledge to never use great apes in its ads again (*1*). Both AAAS (*2*) and CareerBuilder have now made the right move.

CYNTHIA R. SPIESS

Department of Computer Science, Southern Illinois University, Carbondale, IL 62901, USA.

References

1. M. McGraw, "CareerBuilder retires chimpanzee ads following two-year PETA campaign" (23 January 2007); www.nomoremonkeybusiness.com/Newsitem1.asp?id=9447.
2. S. R. Ross *et al.*, *Science* **319**, 1487 (2008).

TECHNICAL COMMENT ABSTRACTS

COMMENT ON "Brain IRS2 Signaling Coordinates Life Span and Nutrient Homeostasis"

Colin Selman, Steven Lingard, David Gems, Linda Partridge, Dominic J. Withers

Taguchi *et al.* (Reports, 20 July 2007, p. 369) reported that mice heterozygote for a null mutation in insulin receptor substrate–2 (Irs2) display a 17% increase in median life span. However, using the same mouse model, we find no evidence for life-span extension and suggest that the findings of Taguchi *et al.* were due to atypical life-span profiles in their study animals.

Full text at www.sciencemag.org/cgi/content/full/320/5879/1012b

RESPONSE TO COMMENT ON "Brain IRS2 Signaling Coordinates Life Span and Nutrient Homeostasis"

Akiko Taguchi and Morris F. White

Differences in reported life span of mice heterozygous for a null allele of the insulin receptor substrate–2 (Irs2) might involve the effects of diet, breeding strategies, and genetic background on insulin-like signaling cascades. A better understanding will emerge from studies focusing on the coordination of nutrient homeostasis and life span by insulin-like signaling in specific peripheral tissues and the central nervous system.

Full text at www.sciencemag.org/cgi/content/full/320/5879/1012c

INTEGRATIVE BIOLOGY

What Systems Biology Is (Not, Yet)

Corey J. Cain,* Debra A. Conte, Marcos E. García-Ojeda, Liza Gómez Daglio, Larry Johnson, Eric H. Lau, Jennifer O. Manilay, Julie Baker Phillips, Nabora Soledad Rogers, Sarah E. Stolberg, Holly F. Swift, and Michael N. Dawson

What is systems biology? Aspects of the field can be traced to the mid-20th century, but its recent growth was sparked by the integration of theoretical and structural molecular biology to answer new questions arising from high-throughput, data-rich, functional genomics. New collaborations, institutes, and at least one new research university (our own) established at the turn of this century embraced systems biology to transform "largely descriptive" biology practiced along disciplinary lines into "a quantitative, predictive" interdisciplinary endeavor (*1*). Almost ten years on, what is systems biology now? As students and faculty drawn together from tissue engineering, molecular and cell biology, physiology, ecology, and evolution into a current topics class of the Quantitative and Systems Biology graduate group at UC Merced, we sought to answer this question to define our course and possibly our futures.

Systems Biology: Philosophical Foundations, a collection of papers arising from a 2005 symposium convened by the Department of Molecular Cell Physiology, Vrije Universiteit, Amsterdam, and the first book on the philosophy of systems biology, was a natural starting point. The editors' "Introduction" describes systems biology as the combination of sciences from "physics to ecology, mathematics to medicine and linguistics to chemistry" and the field's purview as "functional biology." Yet, they see systems biology as largely cell biology. In their view the field primarily studies processes that occur in extant life forms; they also note the importance of research into minimal life ("the smallest unit of life among autonomous cells") and the origin of life.

Surprisingly, given systems biology's supposed broad embrace, the editors explicitly exclude one discipline: evolutionary biology. They explain, quoting Ernst Mayr (*2*), that functional and evolutionary biology are "two largely separate fields which differ greatly in methods, *Fragestellung* [types of questions] and basic concepts." That sentiment is echoed in a chapter on methodologies by Hans Westerhoff and Douglas Kell and again in the editors' "Conclusion" ("systems biology is functional and mechanistic rather than evolutionary biology"), which supposedly summarizes findings of all of the chapters. However, seven of the other 11 chapters discuss the evolution of systems, albeit not always at length. These include thoughtful contributions by William Wimsatt on research programs, Alvaro Moreno on the origin of biological organization, and Evelyn Fox Keller on self-organizing systems.

This discrepancy seems to result from the editors' position that functional systems biologists "may use reasoning derived from evolutionary biology" but ignore the evolution of systems. That claim is at best whimsical, and the editors themselves cite homology of DNA sequences as evidence of the "unity of biochemistry," praise "the successes of … phylogenetics," and foresee synergy with "evo-devo." A more convincing exclusion of evolutionary biology could have been damaging. The mantra that "systems biology tries to understand life as it is now" but need not ask "why-questions" courts the "adaptationist programme" excoriated by Stephen Gould and Richard Lewontin (*3*) three decades ago. Furthermore, the belief that "biological systems live in the absence of evolution" precludes rational treatment of viruses that evolve within the host's body during disease progression (*4*). Fortunately, this is not the current prospectus for the field (*5*–*7*).

Although we initially found *Philosophical Foundations* to be a source of frustration—a common response for scientists attempting their first philosophical texts—the book stimulated self-discovery and eventual insight. Despite our disagreement with some specific conclusions, we appreciated the authors' success at motivating introspection. Because the contributions in the volume are autonomous, individual chapters can be read to usefully complement other systems biology books.

Systems Biology
Philosophical Foundations
Fred C. Boogerd, Frank J. Bruggeman, Jan-Hendrik S. Hofmeyr, and Hans V. Westerhoff, Eds.
Elsevier, Amsterdam, 2007. 360 pp. $120, £68.99, €99.95. ISBN 9780444520852.

An Introduction to Systems Biology
Design Principles of Biological Circuits
by Uri Alon
Chapman and Hall/CRC, Boca Raton, FL, 2006. 317 pp. $54.95, £30.99. ISBN 9781584886426. Mathematical and Computational Biology, vol. 10.

BROWSINGS

Inspired by Nature: Plants. The Building/Botany Connection. Alejandro Bahamón, Patricia Pérez, and Alex Campello. W. W. Norton, New York, 2008. Paper, 192 pp. $35, C$38.50, £19.99. ISBN 9780393732511.
The authors showcase recent buildings in which the study and reinterpretation of plants was used to enhance architectural design. They present examples from around the world that draw on vegetal strategies to create efficient and sustainable forms. Individual chapters cover designs that reflect competition for light or space, control of water, temperature regulation, adaptations to cold or windy settings, and defense mechanisms. The final chapter offers cases—such as the Billboard Building in Tokyo (below)—that adopt aspects of the botanical world for aesthetic or evocative purposes.

CREDIT: FROM *INSPIRED BY NATURE: PLANTS*, COURTESY W. W. NORTON

The reviewers are students and faculty in the Current Topics in Quantitative and Systems Biology class (2007–2008), School of Natural Sciences, University of California, Merced, Post Office Box 2039, Merced, CA 95340, USA.

*Author for correspondence. E-mail: ccain@ucmerced.edu

Uri Alon's *Introduction to Systems Biology*, written as a guide for students, takes a very different approach, one we found more immediately accessible and useful. Alon (a physicist turned molecular biologist at the Weizmann Institute of Science) aims to provide a mathematical framework, to illustrate select design principles, and thus to foster understanding of biological networks. He does so by using introductory equations and logic to capture the essence of natural phenomena whose relevance to biology is illustrated with examples from familiar models: bacterial chemotaxis; the neuronal network of *Caenorhabditis elegans*; *Drosophila* development; and the transcription network, flagellar motor, and Lac operon of the bacterium *Escherichia coli*.

The 12 chapters provide a toolbox for investigating systems in living organisms. They build sequentially to cover three topics: (i) A small subset of all possible networks provides the motifs that are the building blocks of circuits and perform defined information-processing functions in dynamic networks. (ii) Biological circuits are robust to naturally occurring fluctuations in dynamic systems. (iii) Circuits in different systems repeatedly converge on optimal designs. Alon argues that general biological laws have resulted in inherently simple biological networks that employ a few motifs, are robust, and are modular. He is largely convincing that these characteristics are widespread, thus encouraging students of any stage that, using the toolkit outlined, they can develop an intuitive understanding of the general principles.

Although Alon's approach provides a common language potentially scalable from biochemical reactions, through cell-cell signaling, to organismal interactions over evolutionary time, the choice of examples "most familiar to the author" leaves out applications to non-model and population-level systems. Whether it will be practicable in such instances remains to be demonstrated. Along with inexact usage of a few terms—adaptation, fitness, and optimality—such unknowns illustrate the necessity for collaboration if one is to achieve seamless interdisciplinarity. Yet, Alon's approach is technical, not intrinsically functional nor necessarily evolutionary; applied to cellular processes occurring in extant life forms that must be robust through time, it agnostically hints at unification of "understanding of evolved and designed systems."

What is systems biology now? The chapters in *Philosophical Foundations*, their comparison with Alon's approach, and our own experience demonstrate that it is different things to different people. It is perhaps easier and more important to identify what systems biology currently is not. It is not a holistic interdisciplinary endeavor. Systems biology strongly reflects its origins in molecular and cell biology departments, which separated from organismal biology decades ago (*8*). The broad jurisdiction of systems biology is implied by titles of new journals that reference component parts, such as *Molecular Systems Biology*, *Systems and Synthetic Biology*, and *IET Systems Biology*. These, however, perhaps speak louder of what systems biology risks remaining: practiced within traditional disciplinary boundaries. When Linus Pauling was promoted in 1931, he chose "the title of Professor of Chemistry—not theoretical chemistry … not physical chemistry" (*9*). Systems biologists similarly need to choose interdisciplinarity.

The quote from Mayr, employed in *Philosophical Foundations* to exclude the history of life, actually introduces his observation that functional and evolutionary biology "have many points of contact and overlap. Any biologist working in one of these fields must have a knowledge and appreciation of the other field" (*2*). Choosing Mayr's words to eschew evolutionary biology is additionally incongruent because he was among those who led the modern synthesis that linked genetic processes to micro- and macro-evolution and now forms the foundation for contemporary evolutionary biology. That synthesis achieved for organismal biology what systems biology aspires to today. Its far-reaching consequence is epitomized in Theodosius Dobzhansky's famous claim, "nothing makes sense in biology except in the light of evolution" (*10*). Dobzhansky did not demand primacy of evolution over all biology. Likewise, no modern evolutionary biologist could reasonably deny the fundamental contributions made to their discipline since the 1950s by physicists and chemists working on molecular biology. Rather, Dobzhansky was emphasizing that even functionally peculiar biological phenomena—such as the giraffe's recurrent laryngeal nerve and human's forward-articulating knee—can be explained evolutionarily. To Dobzhansky, extant systems and their ancestors were "interdependent and connected by reciprocal feedback relationships" (*10*) retained in modern function. Systems biology's history will, in time, explain what the endeavor becomes. Not by chance alone will it have the good fortune to unify biology.

Phylogeny and Evolution of the Mollusca

Winston F. Ponder and David R. Lindberg, Eds.

University of California Press, Berkeley, 2008. 481 pp. $49.95, £29.95. ISBN 9780520250925.

References

1. A. Agrawal, *Nat. Biotechnol.* **17**, 743 (1999).
2. E. Mayr, *Science* **134**, 1501 (1961).
3. S. J. Gould, R. C. Lewontin, *Proc. R. Soc. London Ser. B* **205**, 581 (1979).
4. B. T. Grenfell *et al.*, *Science* **303**, 327 (2004).
5. M. Medina, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 6630 (2005).
6. N. Kashtan, E. Noor, U. Alon, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 13711 (2007).
7. M. P. H Stumpf *et al.*, *Trends Ecol. Evol.* **22**, 366 (2007).
8. W. J. Bock, *Am. Sci.* **86**, 186 (1998).
9. http://osulibrary.oregonstate.edu/specialcollections/coll/pauling/bond/quotes/allt.html.
10. T. Dobzhansky, *Am. Zool.* **4**, 443 (1964).

10.1126/science.1157405

EVOLUTION

Hard Facts About Soft Animals

Matthias Glaubrecht

Remarkably, the poet Edgar Allan Poe was among the first to comment that a reliable classification of molluscs requires a combined analysis, which in his times meant reconciling a system based on hard shells with evidence from soft body anatomy. In *The Conchologist's First Book* (*1*, *2*), his only book to achieve commercial success during his lifetime, Poe explicitly distinguished between conchology (the study of shells) and malacology (the study of molluscs). Unfortunately, for almost another century conchologists continued to classify molluscs using almost exclusively features of their shells while neglecting information about the soft bodies inside.

With some 200,000 extant species, the Mollusca (snails, clams, and squids) are one of the largest animal phyla, second only to the arthropods (spiders, crabs, insects, and their like). The remarkably rich fossil record of molluscs illuminates their history back into the earliest Cambrian (543 million years ago), and they have since spread into nearly every ecosystem on Earth. The seven or eight classes of living molluscs and two extinct class-rank taxa comprise an array of extremely varied body plans. Molluscs range from minute wormlike animals dwelling among sand grains on the beach to giant squid in the deep sea and from microscopic snails in leaf litter to giant clams in coral reefs. As objects of fascination,

The reviewer is at the Museum of Natural History, Humboldt University, Invalidenstrasse 43, 10115 Berlin, Germany. E-mail: matthias.glaubrecht@museum.hu-berlin.de

function, and food, molluscs play important roles in many cultures and societies. They include many taxa of immense economic significance, such as oysters, scallops, and squids; some bivalves produce precious pearls; some snails carry diseases that infect millions of people, especially in the tropics.

Yet, we still know few hard facts about the evolution and phylogeny of all these soft-bodied animals. For the greater part of the previous century, Johannes Thiele's epochal *Handbuch der Systematischen Weichtierkunde* (*3*), published between 1929 and 1931, provided the standard in molluscan systematics. For his classification Thiele, a malacologist at the Natural History Museum in Berlin, evaluated characteristics from shell, radula, and anatomy in a synthetical, albeit precladistic, manner. Although he sometimes erred, as for example when he discarded aplacophoran molluscs as worms belonging to the phylum Annelida, Thiele's systematization was a masterpiece. By the time his handbook was finally translated into English and republished in 1992 (*4*), a substantial volume of new data from morphology, molecular genetics, and paleontology had accumulated, already too diverse for any one malacologist to master.

New tools, from molecular techniques to fluorescence-coupled antibody staining and confocal laser scanning microscopy, in concert with new approaches, such as computer-assisted cladistics, that allow the recurrent testing of phylogenetic hypotheses under various models and assumptions have recently generated a renewed interest in reconstructing animals' evolutionary history. Consequently, within the past two decades our understanding of molluscan phylogeny has undergone a remarkable transformation that is about to fundamentally change the classification of the phylum.

Phylogeny and Evolution of the Mollusca provides an updated collation badly needed at a time when seemingly every new phylogenetic tree calculated using another partial gene fragment from any more-or-less randomly chosen set of taxa is considered to merit its own publication. The volume, compiled by Winston Ponder (a recently retired researcher at the Australian Museum in Sydney) and Dave Lindberg (University of California in Berkeley), two of the most distinguished malacologists of our time, grew out of a symposium organized by the editors during the 2004 World Congress of Malacology in Perth. The symposium aimed to review the available evidence from which we hypothesize molluscan phylogeny, and the book brings together what we know so far from disparate areas such as the study of ultrastructure of sperm, chemosensory organs, and many other morphological structures as well as sequence data from mitochondrial and nuclear gene fragments and multigene analyses. In doing so, it offers a phylogenetic platform from which malacology can proceed into the genomic era.

Most of the contributors are authorities on the topics they cover. The first three chapters offer a general introduction to molluscs, review recent work on the phylogenetic position of Mollusca and relations among its higher taxa, and consider putative Early Cambrian relatives such as *Kimberella*, *Wiwaxia*, and the halkieriids. The following 12 chapters provide a taxon-by-taxon synopsis of the features and evolution of each of the major groups, from wormlike marine Solenogastres and Caudofoveata, through the shell-bearing uni- and bivalved forms, to the "naked" air-breathing land pulmonates. Each of these chapters reviews the data in a systematic way, with a shared format that helps readers quickly orient themselves. A few chapters also provide original data and new analyses—for example, a combined morphological and molecular phylogenetic study on gastropods (which, with about 150,000 extant species, comprise 80% of all molluscs).

The penultimate chapter approaches molluscs from an evolutionary-developmental viewpoint, revealing how little we still know in this field. The book ends with an overview of molluscan evolutionary genomics, including a useful survey of available data from high-throughput DNA sequencing. That chapter underscores our present position at the dawn of a new synthesis of biosystematics and evolutionary biology, paralleling that of the early 20th century when the fledgling field of genetics identified the underlying basis for anagenesis and cladogenesis.

Readers may find that they disagree with certain authors' approaches and claims here and there (e.g., the concept of reconstructing a hypothetical ancestral mollusc). I am unhappy with the editors' and, under their guidance, most authors' uncritical use of the reappearing term "adaptive radiation." In contrast to Henry Fairfield Osborn's traditional concept of a fan-shaped diversification pattern, that is now much more strictly (and, to my taste, adequately) defined by evolutionary ecologists (*5*). For some topics, contributors could have provided a more comprehensive coverage of the literature.

However, these minor points aside, the volume establishes how much malacology can contribute to a future synthesis reconciling morphology with molecules. The book makes exciting reading, all the more so because in providing overviews of work in progress it reflects divergent opinions. Recurrent sections on our gaps in knowledge and the authors' recommendations for future studies reveal how much and what remains to be done. *Phylogeny and Evolution of the Mollusca* cannot be the definitive summary, but it constitutes a large step forward. The review provided by Ponder, Lindberg, and their colleagues should persuade a new generation of malacologists to follow Edgar Allan Poe's insight.

Two of the many morphologies of molluscs. The snail (left), an undescribed *Tylomelania*, is one of about 40 members of a lacustrine species flock endemic to Sulawesi. Solenogastres, such as *Falcidens halanychi* (right), are mostly minute, worm-shaped inhabitants of deepwater marine habitats.

CREDITS: (LEFT) CHRIS LUNAU/P.WHNK CRUSTA 10.DE; (RIGHT) C. F. CHRISTOFFER SCHANDER

References and Notes

1. E. A. Poe, *The Conchologist's First Book: Or, A System of Testaceous Malacology, Arranged Expressly for the Use of Schools...* (Haswell, Barrington, and Haswell, Philadelphia, 1839).
2. Although published under Poe's name, the book was essentially a less-expensive edition of a 1838 work by Thomas Wyatt—who plagiarized most of the text from naturalist Thomas Brown's *The Conchologist's Textbook* (Glasgow, 1837). Poe wrote the preface and introduction.
3. J. Thiele, *Handbuch der Systematischen Weichtierkunde* (G. Fischer, Jena, 1929–1931).
4. J. Thiele, *Handbook of Systematic Malacology*, R. Bieler, P. M. Mikkelsen, Eds., J. S. Bhatti, Transl. (Smithsonian Institution Libraries and National Science Foundation, Washington, DC, 1992–1998).
5. D. Schluter, *The Ecology of Adaptive Radiation* (Oxford Univ. Press, Oxford, 2000); reviewed by A. Meyer, *Science* **294**, 64 (2001).

10.1126/science.1158066

PUBLIC HEALTH

Public-Private Partnerships and Scientific Imperialism

Developed industrial nations are not taking their research and development colleagues of the developing world as equal partners in fighting infectious disease.

T. J. Tucker[1,2*] and M. W. Makgoba[3*]

Diseases that dominate the health of most African populations, such as AIDS, malaria, and tuberculosis, have always received a small proportion of the global financial support that is available for medical science and health interventions. They are thus often referred to as "neglected diseases." This is reflected in the number of new drugs developed in recent decades. Of the 1393 new drugs approved in the 25 years before 2000, only 13 were specifically indicated for tropical diseases (*1*). Health-related public-private partnership organizations (PPPOs)—nongovernmental organizations committed to a "hybrid" set of activities, including focused product development, advocacy, and capacity building—have been important in generating resources and novel products (*2*). However, it is argued that they have not been able to change the prevailing imperialist paradigm or to involve African researchers as equally empowered participants and leaders.

The pursuit of novel scientific interventions for AIDS, malaria, and tuberculosis has been supported for decades by the traditional public-sector research funding bodies, such as the U.S. National Institutes of Health, the U.K. Medical Research Council, the Agence nationale de recherches sur le sida, and the European and Developing Country Trial Partnership, with additional contributions from the private sector and charitable bodies. These major public-sector funding bodies are located in developed countries and, although the situation is changing, direct access to funding from almost all of these agencies has historically been very difficult (or legally impossible) for researchers from developing countries. As a result, even where developing-country researchers have received research funds from such agencies, most of the funding has been channeled through the host country institutions. This creates a dependency relationship, as well as multiple bureaucratic hierarchies in administering such grants. Although certain developing countries have research bodies that fund research into the neglected diseases, the quantity of funding is generally very small.

The last decade has seen the establishment of large health-related PPPOs, each of which aims to address specific scientific or other hurdles relating to one or more of the neglected diseases. This includes infrastructure and capacity development in funded sites. Each of these PPPOs has a different mandate, which varies from preventive health interventions such as vaccines and microbicides to novel drugs and diagnostic tests. The PPPOs fund product development work, and then own or co-own the products that emerge, with the intention of facilitating the uptake of these products into global health care environments.

The establishment of health-related PPPOs received an enthusiastic early endorsement from African stakeholders, because they were created with the explicit mandate of making a distinct break from the previous developed-country, dependency-based, research paradigm. A good proportion of these PPPOs have achieved significant progress over time, and

PPPO Demographics

Name	Head office	CEO Gender	CEO Nationality	Percent from Africa on Board of directors	Percent from Africa on Advisory board(s)
International AIDS Vaccine Initiative	New York, NY, USA	Male	American*	7.7	14.3
Global Alliance for TB Drug Development	New York, NY, USA	Male	French	23.0	8.3
Malaria Medicines Venture	Geneva, Switzerland	Male	British	20.0	5.5
Aeras Global TB Vaccine Foundation	Washington, DC, USA	Male	American	0.0	3.4
Program for Appropriate Technology in Health	Seattle, WA, USA	Male	American	23.1	0.0†
International Partnership for Microbicides	Silver Spring, MD, USA	Female	American	10.0	15.4
Foundation for Innovative New Diagnostics	Geneva, Switzerland	Male	Italian	25.0	Not given
Institute for OneWorld Health	San Francisco, CA, USA	Female	American	0.0	8.7
Global HIV Vaccine Enterprise	New York, NY, USA	Male	Canadian	Pending	17.6‡

*All Americans are from the USA. †Leadership council. ‡Coordinating committee.

Demographics of some major PPPOs operating in the field of neglected diseases.

[1]Scientific Medical Research (Proprietary) Ltd., Mandela-Rhodes Building, St. Georges Mall, Cape Town, South Africa; e-mail: tim@smr.co.za (T.J.T.). [2]Division of Medical Virology, Department of Clinical Laboratory Sciences, University of Cape Town, Cape Town, South Africa. [3]University of KwaZulu-Natal, University Road, Westville, Durban, South Africa; e-mail: makgoba@ ukzn.ac.za (M.W.M.).

some have grown from small organizations into large bodies with hundreds of employees performing excellent work within a complex global network. This should be applauded, because the scientific and logistic hurdles faced by PPPOs are substantial. PPPOs have managed to access significant new financial resources to support their scientific investigations, through the staggering generosity of charitable foundations such as the Bill and Melinda Gates Foundation and others, as well as monies from private and public funds. The competition among research groups for these funds is intense.

However, on examining the Internet-available data, we question whether the current PPPO paradigm has, in fact, perpetuated research disparities and power inequities and possibly accentuated the dependency relationship of developing-country researchers rather than contributing to correcting the disparity. Data show that the notion that PPPOs are global organizations with equal (or at least significant) representation from all regions of the world is false. Although all health-related PPPOs focus on neglected African diseases, they are distinctly first-world entities with differing levels of outreach to developing countries. Every major PPPO is headquartered in the United States or Europe, and most are based in the United States (see table, page 1016). As a result, almost all monies raised by PPPOs are channeled through first-world head offices, and any decisions made regarding how these are spent in developing countries are made by the CEOs, together with their senior staff and boards (statutory and advisory). A great majority of CEOs are male, all are Caucasian, and all are residents of either the United States or Europe. Not one "global" PPPO is led by a person who is a developing-country national, and not one resides within one of the developing countries severely affected by neglected infectious diseases.

Only a very small proportion of senior staff at the executive director level has non-U.S. or non-European origins. The PPPO boards of directors, who have ultimate responsibility for the organization's actions, show similar trends, with Africans generally making up a small proportion of statutory board membership (0 to 20%). The advisory boards (or leadership council) generally have better representation from Africa than statutory boards, but representation remains low (0 to 24%). The main advisory boards tend to have fewer representatives from Africa than the less influential subadvisory committees or boards. In addition, the main advisory boards of PPPOs mostly have no people representing "on-the-ground" communities in Africa, despite the fact that this input is critical if large studies are to have cultural sensitivity in resource-poor environments.

There have been impressive disbursements to developing countries, with PPPOs such as the Program for Appropriate Technology in Health, the International AIDS Vaccine Initiative, and others investing millions in capacity development, infrastructure, and research and development, activities that have the potential to benefit Africa. However, the true benefit of these has yet to be assessed.

However well-intentioned these individuals and organizations are, the traditional imperialist power dynamics remain, in which African researchers have very limited executive decision-making ability within PPPOs, and Africans are only able to access resources that (predominantly) non-Africans decide are appropriate. It is the dominant neocolonial structure of funding and operation that characterizes all PPPOs irrespective of grant sizes that is of major concern.

There is no doubt that African research groups and communities play a meaningful role in PPPO work. They perform critical clinical work associated with product development, which often cannot be performed in the developed countries. However, the nature, scope, and budget of that work are almost always ultimately decided upon by the PPPO head offices, not the African researchers, and thus executive decision-making remains outside Africa.

Before it is thought that this paper argues in favor of reducing the involvement of those from North America and Europe, let us clarify. It is abundantly clear that there is limited capacity in Africa and that many of the hurdles faced by the PPPOs are scientific and not logistic. It is also true that the overwhelming majority of scientific resources in the world are in North America and Europe. Africans know this. Thus, we need to argue vociferously for increased involvement of all communities to solve the problems we face.

The African research leadership community must also ask whether they are not, in part, responsible for this situation as well. African countries have been notorious for corruption, and much donor funding remains unaccounted for (*3*). Yet, many institutions appear not to be taking adequate steps to counteract such corrupt practices. We have played senior roles in the South African AIDS Vaccine Initiative (SAAVI) (*4*), and we believe that it is a positive example of financial accountability. It is likely that there will be far higher levels of willingness shown to fund African PPPO efforts if more attention is paid to issues of corruption. In addition, African states have not contributed sufficiently to creating career structures for clinicians and scientists within the African institutions, so there is a relative lack of available capacity to build PPPOs in Africa. If more of these career paths were available to African scientists, fewer would be drawn away from the continent, and thus, the establishment of PPPOs in Africa would be easier.

African states should themselves be investing in health-related PPPOs, as too few examples of successful African-funded science programs exist. Global financiers would likely be more willing to fund African-based initiatives if individual African countries or the African Union structures led the way in establishing and funding PPPOs to address important health issues. Having greater African leadership in this is critical.

African scientists and clinicians also need to address power imbalances when establishing their own individual relationships with the large PPPOs. Although many African researchers are desperate for any funds that become available, others who are relatively well resourced appear to perpetuate the power imbalances by accepting contractual relationships with PPPOs that are continuing the old dependency paradigms. African researchers themselves need to take a more proactive stance on these matters and to ensure that their relations with the large PPPOs and funders are better balanced.

Africa's relative lack of scientific capacity accentuates the need to increase Africa's meaningful role in PPPOs, lest PPPOs themselves increase the scientific capacity disparities, power inequities, and scientific deficiencies that already exist. It is imperative that those who fund and control PPPOs change the way in which PPPOs operate, including their structures, personnel, and systems, so that Africans will not remain relatively disempowered participants. Africa should not dominate PPPO structures. However, the current PPPO paradigm is fundamentally neocolonial in structure and operation, and this needs to be revisited and addressed.

References and Notes

1. P. Trouiller *et al.*, *Trop. Med. Int. Health* **6**, 945 (2001).
2. J. Chataway *et al.*, *Eur. J. Dev. Res.* **19**, 100 (2007).
3. L. Eaton, *BMJ* **331**, 718 (2005).
4. SAAVI is housed within the South African Medical Research Council (MRC), a governmental statutory body that accounts to the public through reporting to the parliament of the Republic of South Africa, and its finances are audited annually. The annual financial audits of the MRC from 1998 to 2008 are available at www.mrc.ac.za.

10.1126/science.1156720

ENGINEERING

Wi-Fi-Fo-Fum

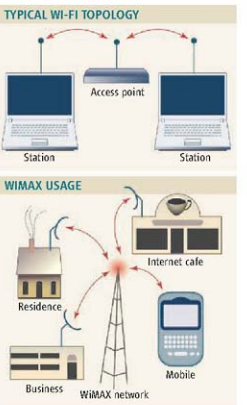
A guide to what actually happens when you connect your laptop with Wi-Fi and WiMAX networks.

Robert Morrow

For more than a decade, the brawny giant in wireless communication has been the Wi-Fi implementation. So-called "hot spots" abound, saturating homes, airports, hotels, and restaurants. With an indoor range of 10 to 50 m, depending on data rates and obstacles (*1*), one Wi-Fi base station can serve users throughout a small building (see the first figure, top panel). Larger areas can be supported via multiple base stations, and even seamless roaming is feasible for those with the motor skills to walk and surf at the same time. The latest version of Wi-Fi equipment, based on the draft version of the IEEE 802.11n standard, promises to deliver raw data rates as high as 600 megabits per second (Mb/s), which is sufficient to carry several simultaneous high-definition television signals. Efforts to expand the useful range of Wi-Fi even farther have run into obstacles of cost and interference from other radio-frequency devices. These difficulties have been addressed with a new standard, WiMAX, which operates in the licensed radio band at higher transmit power (see the first figure, bottom panel).

Such widespread use did not come overnight. Begun in 1990 by the Institute of Electrical and Electronics Engineers (IEEE) as Project 802.11 (*2*), the Wi-Fi specification has changed substantially as new theory and implementation methods evolved. The first 802.11 standard, published in 1997, supported raw data rates of 1 and 2 Mb/s. International acceptance of the standard was aided by operating in the (mostly) worldwide 2.4-gigahertz (GHz) license-free radio band. The radios were expensive and power-hungry, and devices from different manufacturers often refused to communicate with each other. Despite these restrictions, corporate users in particular embraced 802.11 as a way to connect computers together without cables.

A banner year for 802.11 evolution was 1999, out of which emerged three major improvements. The first, called 802.11b, supported faster data rates of 5.5 and 11 Mb/s. The second was the formation of the Wi-Fi Alliance by several vendors to provide a means for interoperability testing and certification. Earning the Wi-Fi logo guaranteed that the product could communicate with other certified products, regardless of who made them. The third improvement was the release of 802.11a, which specified an entirely new process of digital radio communication called orthogonal frequency division multiplexing, or OFDM (*3*).



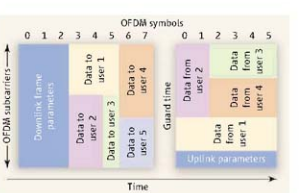
Wi-Fi and WiMAX landscapes. The lower-power Wi-Fi implementation can readily connect devices within a building, whereas WiMAX can connect devices over greater distances.

This multicarrier modulation method uses 48 data subcarriers and four pilot subcarriers, all spaced as closely as possible in frequency within the band without causing mutual interference. The pilot subcarriers act as references for changes in transmission properties in different operating environments. Each subcarrier holds its own independent data stream, which allows aggregate data rates of 6 to 54 Mb/s in the 5-GHz license-free communication bands. Data are transmitted in 4-μs windows called symbols; by combining amplitude and phase information, multiple bits can be transmitted in a given symbol (*4*).

The 802.11a standard was slow to be adopted, however, because inexpensive 5-GHz radios were difficult to make, the 5-GHz band was not available worldwide, and the digital signal processing (DSP) to build the OFDM signals overwhelmed low-cost late 20th-century technology. In June 2003, with the release of the 802.11g standard, all of these impediments were addressed. With a change in government wireless regulations in several countries, OFDM was now allowed in the 2.4-GHz band—which, by operating at lower frequency, allowed the components to be larger and hence cheaper to manufacture. Also, DSP became sufficiently fast and sophisticated to build a new OFDM signal every 4 μs, as required by the standard.

Giants in technology, to remain giants, must always be on the move, and 802.11 is no exception. The latest enhancement, 802.11n, is set to become an official IEEE standard in late 2008. Taking another leap upward in sophistication, 802.11n still uses OFDM but now allows up to four transmit antennas, each radiating its own independent data. Given sufficient DSP, these antennas can also use beam forming; that is, they can steer their signals in the direction of the desired receiver.

In a similar way, multiple antennas at the receiver can electronically aim their sensitive main lobe toward the desired signal source, combine reflected copies (multipath components) of the signal, and even turn a deaf ear toward interfering transmissions. Both the 2.4- and 5-GHz license-free bands are sup-



WiMAX management of multiple users. Mobile WiMAX processes signals in 102.9-μs intervals, or "symbols," and can allocate groups of subcarriers to different users during both downlink and uplink frames.

CREDIT: ADAPTED FROM ROBERT MORROW

Morrow Tech Services, 6976 Kempton Road, Centerville, IN 47330, USA. E-mail: rkmorrow@ieee.org

ported. All of this technology can now be had for just a fraction of the cost and physical size of basic 802.11 devices that were prevalent at the beginning of this decade.

As versatile as Wi-Fi may be, however, it cannot easily be shoehorned into applications that are able to service large regions of fixed, portable, and mobile users at low cost and ultralow power. Indeed, several municipalities abandoned earlier plans to install citywide Wi-Fi coverage when they became aware of the cost and complexity of placing thousands of base stations throughout the community.

In an effort to provide long-range, high-speed data access, at least for nonmobile users, the Fixed WiMAX standard (IEEE 802.16-2004) was developed. For short-range wireless audio connections to a cell phone, Bluetooth has a near-monopoly, and ZigBee (who dreams up these names?) takes aim at simple, low-cost security and control applications (*5*).

Providing high-speed data services to highly mobile users is the next great challenge in terrestrial wireless engineering. Solutions range from various third-generation (3G) cellular-based proposals to an IEEE standard called 802.16e-2005 with the commercial name Mobile WiMAX. Taking a page out of the Wi-Fi standard, WiMAX has much in common with 802.11n: OFDM (now with up to 2048 subcarriers), multiple antennas carrying independent data streams, and beam forming (*6*).

Perhaps the greatest advantage WiMAX has over Wi-Fi is that it operates in various licensed frequency bands, thus eliminating the cacophony of interference prevalent on the unlicensed bands. In addition, as a licensed service, regulations allow a WiMAX transmitter to have much higher power, easily expanding its useful range to 1 km or more, or at least 20 times the reach of a Wi-Fi base station; an entire city can be serviced with a few dozen base stations.

The flip side of such a large coverage area in a mobile user environment is the chance that a large number of users may at times converge on a single WiMAX base station, which will require sophisticated resource allocation methods. WiMAX provides such sophistication by assigning different OFDM subcarrier groups to different users. In this way, a single WiMAX transmission frame can carry data to several users at once, either on the uplink (user → base station) or downlink (base station → user) (see the second figure). Furthermore, if throughput becomes too low, a WiMAX base station can direct some of its users to connect to another base station.

It is likely that Wi-Fi radios will remain the dominant player in indoor areas with a relatively high user density. WiMAX may very well become the system of choice where user density is lower and larger coverage areas are needed. Thus, the best way to stay fully connected may require having access to both technologies. Fortunately, dual-mode WiMAX and Wi-Fi devices are already available.

References and Notes

1. T. Rappaport, *Wireless Communications: Principles and Practice* (Prentice-Hall, Upper Saddle River, NJ, 1996).
2. C. Links, W. Diepstraten, V. Hayes, *BYTE* (May 1994), pp. 99–108.
3. J. LaRocca, R. LaRocca, *802.11 Demystified* (McGraw-Hill Telecom, New York, 2002).
4. For more on OFDM operation, see www.complextoreal.com/chapters/ofdm2.pdf.
5. Web sites for the commercial wireless products in this article: www.wi-fi.org, www.wimaxforum.org, www.bluetooth.org, www.zigbee.org.
6. J. Andrews, A. Ghosh, R. Muhamed, *Fundamentals of WiMAX* (Prentice-Hall, Upper Saddle River, NJ, 2007).
7. The terms Wi-Fi, 802.11, Bluetooth, and ZigBee are registered trademarks of the Wi-Fi Alliance, IEEE, the Bluetooth Special Interest Group, and the ZigBee Alliance, respectively.

10.1126/science.1153827

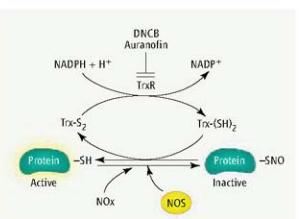
BIOCHEMISTRY

SNO Removal

Reduced thioredoxin denitrosylates caspase-3, a key protein involved in cell death.

Arne Holmgren

More than 100 proteins become S-nitrosylated in cells (*1*). Nitric oxide acts as a mild oxidant (*1*, *2*) that reacts with a cysteine residue's thiol group (thiolate) to form an S-nitrosothiol (SNO). This modification affects most classes of proteins and alters protein function. Thus, in the context of redox-based regulation of cell signaling, it is a highly dynamic posttranslational regulatory mechanism of physiological and pathophysiological importance. However, the mechanisms that remove nitric oxide (denitrosylation) have largely remained uncharacterized. On page 1050 of this issue (*3*), Benhar *et al.* show that thioredoxin and thioredoxin reductase in the cytosol and mitochondria, respectively, constitute a major physiological denitrosylation mechanism that controls programmed cell death (apoptosis) in response to particular stimuli.



Dynamic equilibrium. Formation of S-nitrosothiol (SNO) on a protein thiol (SH) by nitric oxide donors (NOx) or nitric oxide synthase (NOS) changes a protein's activity. Removal of SNO by reduced thioredoxin [Trx-$(SH)_2$] restores the thiol. Oxidized thioredoxin (Trx-S_2) is reduced by nicotinamide adenine dinucleotide phosphate (NADPH), catalyzed by thioredoxin reductase (TrxR). dinitrochlorobenzene (DNCB) and auranofin are inhibitors of TrxR.

A subpopulation of caspase-3 in the mitochondria is constitutively S-nitrosylated and, as such, is inhibited. Stimulation of the cell surface "death" receptor Fas causes the denitrosylation of this subpopulation of caspases, leading to apoptosis (*4*). However, the mechanism of denitrosylation has not been clear.

Benhar *et al.* developed an assay based on reactivation of nitrosylated, inactive caspase-3 and denitrosylating activity that was purified from cell extracts. They found that thioredoxin and thioredoxin reductase together act as a principal denitrosylation mechanism. Thioredoxin, a conserved dithiol protein with a Cys-Gly-Pro-Cys active site (*5*), reduces

The enzyme caspase-3 cleaves several target proteins and is the major executioner of apoptosis. This protease is inhibited by nitrosylation of a cysteine residue in its active site.

Medical Nobel Institute for Biochemistry, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, SE-171 77 Stockholm, Sweden. E-mail: arne.holmgren@ki.se

protein disulfides by a cysteine thiol-disulfide exchange mechanism. The active-site disulfide in oxidized thioredoxin is itself reduced by thioredoxin reductase using electrons from nicotinamide adenine dinucleotide phosphate. Human and other mammalian thioredoxin reductases contain an active-site selenolthiol that mediates this reduction (*6*).

By inhibiting thioredoxin reductase in cultured T cells with the compounds dinitrochlorobenzene and auranofin (both target the active-site selenol in the reduced enzyme), Benhar *et al.* show that thioredoxin and thioredoxin reductase are the physiologically active principle. This was confirmed using RNA interference to reduce the expression of thioredoxin and thioredoxin reductase in cultured T cells. The authors also examined human macrophages treated with lipopolysaccharide and interferon-α, which induce nitric oxide production. They show that this treatment, combined with inhibition of thioredoxin reductase by auranofin, elicits robust S-nitrosylation of caspase-3. In human cells, thioredoxin 1 and 2 are the cytosol-nuclear and mitochondrial isoforms, respectively, and the products of two separate genes; the same is true for thioredoxin reductase. These results show that thioredoxin is the major denitrosylating agent in this cell type, and together with thioredoxin reductase, controls the amount of basal S-nitrosylation of caspases in response to Fas activation (*4*). Thioredoxin reductase may be rate-limiting for the overall effects.

Originally identified as a hydrogen donor for ribonucleotide reductase, thioredoxins have a long evolutionary history (*5*, *6*). Thioredoxin is a major reductant for keeping sulfhydryl groups on intracellular proteins in a reduced state, thus controlling redox states and signaling in the cytosol and nucleus. Thioredoxin also controls the redox state of thiols that are oxidized by hydrogen peroxide. As part of the cell's antioxidant defense, thioredoxin donates electrons to peroxiredoxins, enzymes that catalyze the breakdown of hydrogen peroxide. The enrichment of thioredoxin 1 at the plasma membrane (*6*), where enzymes that generate nitric oxide (nitric oxide synthases) are also concentrated, suggests that signaling by nitrosylation may be a localized phenomenon. Reactions between hydrogen peroxide, peroxynitrite, or nitric oxide donors and protein cysteine thiols may give rise to sulfenic acid, which is reduced by thioredoxin (*7*).

Previous results identified thioredoxin and thioredoxin reductase in denitrosylating the low molecular weight nitrosylating agent S-nitrosoglutathione in vitro (*8*) and proteins (including caspase-3) in human liver cells (*9*). Benhar *et al.* extend these findings by showing that both thioredoxin 2 in the mitochondria and thioredoxin 1 in the cytosol and nucleus are denitrosylases in the physiological context of apoptosis. The study also indicates the presence of a dynamic equilibrium, whereby nitrosylation and denitrosylation reactions occur (see the figure). The results are important because they identify a denitrosylase pathway in the cell and widen the range of thioredoxin functions to the control of apoptosis. Studying the turnover of thioredoxin and thioredoxin reductase in the context of nitrosylating and denitrosylating reactions should be a rich playing field for determining thiol redox control of membrane receptors, ion channels, and transcription factors (*6*, *7*).

The study by Benhar *et al.* also suggests that the effect of thioredoxin and thioredoxin reductase inhibitors as cancer treatments to control cell viability (*10*) may involve previously unknown effects on S-nitrosylation. Reduced thioredoxin is bound to proteins such as thioredoxin-interacting protein or apoptosis signaling kinase (*7*). These interactions should now be considered in relation to nitrosylation. Human cytosolic thioredoxin contains three structural cysteine residues that control thioredoxin activity (*6*) and potentially control transnitrosylation reactions (*11*). Furthermore, thioredoxin-dependent control of nitric oxide synthase may involve denitrosylation of the synthase (which would stimulate nitric oxide production) (*1*). We are only beginning to understand the links among thioredoxin, protein nitrosylation, and redox-controlled signaling (*3*, *9*, *10*).

References and Notes

1. D. T. Hess *et al.*, *Nat. Rev. Mol. Cell Biol.* **6**, 150 (2005).
2. T. Finkel, *Curr. Opin. Cell. Biol.* **15**, 247 (2003).
3. M. Benhar *et al.*, *Science* **320**, 1050 (2008).
4. J. B. Mannick *et al.*, *Science* **284**, 651 (1999).
5. A. Holmgren, *Annu. Rev. Biochem.* **54**, 237 (1985).
6. C. H. Lillig, A. Holmgren, *Antioxid. Redox. Signal.* **9**, 25 (2007).
7. L. M. Landino, *Methods Enzymol.* **440**, 95 (2008).
8. D. Nikitovic, A. Holmgren, *J. Biol. Chem.* **271**, 19180 (1996).
9. R. Sengupta *et al.*, *Biochemistry* **46**, 8472 (2007).
10. E. S. J. Arner, A. Holmgren, *Sem. Cancer Biol.* **16**, 420 (2006).
11. D. A. Mitchell *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 11609 (2007).

12. Supported by grants from the Swedish Cancer Society, Swedish Research Council, and the K. A. Wallenberg Foundation.

10.1126/science.1159246

OCEAN SCIENCE

Marine Calcifiers in a High-CO_2 Ocean

Victoria J. Fabry

New results show that the response of marine organisms to ocean acidification varies both within and between species.

The oceans have taken up about one-third of the total carbon dioxide (CO_2) released into the atmosphere by human activities over the past 200 years (*1*). This addition of CO_2 to the surface ocean changes seawater chemistry, resulting in a decrease in pH and carbonate ion concentration, and an increase in the concentrations of bicarbonate ion and hydrogen ion. Ocean absorption of anthropogenic CO_2 also reduces the saturation state of seawater with respect to calcite and aragonite, two common types of calcium carbonate secreted by marine biota. Experiments with calcareous organisms indicate that calcification is strongly dependent on the carbonate saturation state of seawater (*2*–*6*), which suggests that ocean acidification will adversely impact calcifying taxa. Evidence is now accumulating that the acidifying effects of CO_2 on seawater may have diverse consequences for marine calcifiers.

Most studies on the effects of ocean acidification on calcification have focused on warm-water coral reef species (*2*–*4*). Only six species of living, planktonic calcifying organisms have been investigated—four coccolithophore species and two foraminiferan species. Similar to coral reef species (*4*) and several other taxa (*2*, *3*), two bloom-forming coccolithophores (including *Emiliania huxleyi*; see the first figure) (*5*, *6*), showed reduced calcification rates in response to elevated partial pressure of CO_2 (pCO_2) in seawater. However, two other coccolithophore

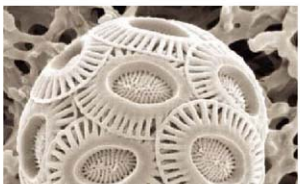
Department of Biological Sciences, California State University San Marcos, San Marcos, CA 92096, USA. E-mail: fabry@csusm.edu

species, which are important in terms of calcium carbonate export to sediments, did not conform to this pattern (*7*). Calcification in one species did not significantly change when seawater pCO_2 varied from 150 to 915 μatm (see the second figure, bottom right). In the other species, calcification rates measured over a similar range of pCO_2 levels resulted in a nonlinear curve (see the second figure, bottom left), with maximum calcification occurring at pCO_2 values that correspond to that of the present surface ocean.

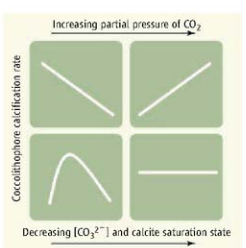
Recently, Iglesias-Rodríguez *et al.* (*8*) confirmed that the response of calcifying organisms to CO_2-induced changes in seawater chemistry is not uniform and may be much more complex than previously thought. When the authors grew laboratory cultures of *E. huxleyi* at 750 parts per million by volume (ppmv) CO_2—almost twice the average value of today's surface ocean—rates of calcification doubled compared to calcification rates of cells grown at 280 ppmv (see the second figure, top right). This is in marked contrast to earlier work with laboratory cultures and field populations of *E. huxleyi* (*5*, *6*), which reported decreases in calcification rates under high-CO_2 conditions (see the second figure, top left).

The authors suggest that one reason their results differ from previous work with this species is that they manipulated the inorganic carbon system of seawater by bubbling with air containing different concentrations of CO_2. In contrast, several previous studies used hydrochloric acid to achieve target pCO_2 values (*5*, *6*). The former method more closely mimics ocean acidification because it increases the concentration of total dissolved CO_2, whereas adjusting the carbonate chemistry with acid holds the concentration of total dissolved CO_2 constant. Moreover, compared to the acid technique, the gas bubbling method results in a larger increase in the bicarbonate ion concentration, thought to be the substrate for calcification in *E. huxleyi* (*9*). Several earlier laboratory and mesocosm experiments bubbled with CO_2-enriched air, yet those studies reported reduced calcification rates in *E. huxleyi* in response to elevated pCO_2 (*6*). Part of the discrepancy in results may be due to other differences in experimental conditions. Factors such as nutrients, light, trace metals, temperature, and salinity, as well as their interactive effects with CO_2, can affect coccolithophore calcification in ways that are just beginning to be understood (*6*).

An alternative hypothesis that may explain the divergent results is that the traditional view



Scanning electron micrograph of the coccolithophore *Emiliania huxleyi*.



Calcification response patterns of four coccolithophore species to elevated CO_2. (Top left) Decreased calcification (*5*, *6*); **(top right)** increased calcification (*8*); **(bottom left)** nonlinear response (*7*); **(bottom right)** no change (*7*). Similar response patterns have been identified in benthic organisms (*12*).

of *E. huxleyi* as a single species is flawed. Increasingly, evidence suggests that this cosmopolitan species is composed of many varieties with genetically and morphologically distinct properties (*10*, *11*), which may account for the physiological variability observed in *E. huxleyi* with respect to calcification and other processes (*9*). Indeed, the high genetic diversity of *E. huxleyi* may confer ecophysiological plasticity that will allow it to persist and even flourish in the face of ongoing ocean acidification and global warming.

A taxonomically varied response in calcification to CO_2-induced ocean acidification is not limited to coccolithophores. Ries *et al.* (*12*) conducted experiments with a variety of calcareous benthic invertebrates and macroalgae over a range of carbonate ion concentrations and aragonite saturation states. Some organisms had increased rates of calcification in response to elevated pCO_2, while others showed decreased calcification rates, and still others displayed a nonlinear, parabolic response as a function of carbonate saturation state. Even closely related organisms—for example, tropical and temperate sea urchins—showed different responses.

Calcification mechanisms in calcium carbonate–secreting organisms are not sufficiently understood to explain the species-specific differences observed in CO_2-perturbation experiments. The energy dependence of calcification and mechanisms to prevent dissolution implies that there are physiological limitations such that net dissolution will exceed net calcification when CO_2 is raised above a threshold value. The results of Iglesias-Rodríguez *et al.* (*8*) and Ries *et al.* (*12*) suggest that such "tipping points" may vary considerably among calcifying species. A differential response of calcifiers to elevated CO_2 conditions may result in competitive advantages that could drive the reorganization of many planktonic and benthic ecosystems, which in turn could have significant ecological and biogeochemical implications.

Currently, persistent aragonite undersaturation of surface seawater in high latitudes is projected to occur as early as 2050 (*13*). Recent observations of seasonal aragonite undersaturation of surface waters in temperate, nearshore regions, however, suggest that some calcareous organisms may already be experiencing substantial, transient changes in seawater CO_2 chemistry (*14*, *15*). Given the importance of coastal and high-latitude regions to fisheries and other ecosystem services, a comprehensive understanding of the impacts of ocean acidification is urgently needed.

Future work should investigate additional species and life stages of calcifiers, examine the possible interactions of increasing CO_2 with other environmental variables, and explore the capacity of organisms to adapt to projected changes. Equally important is the development of internationally agreed-upon, standardized protocols for the control of seawater composition in manipulative experiments, as well as for the measurement of calcification rates. Without such protocols, it will continue to be difficult to compare results among laboratories, across taxa, between regions, or over time.

The diverse pattern of poorly understood biotic responses to ocean acidification found thus far makes it problematic to reliably predict the ecological and biogeochemical changes that will result from continued oceanic uptake of anthropogenic CO_2. As

PHOTO: GEOFF JENDEN KOUNOUNARTURAL HISTORY MUSEUM, LONDON

atmospheric CO_2 levels continue to rise, we are embarking on a global experiment with as yet uncertain long-term consequences for many marine calcifiers.

References

1. C. L. Sabine *et al.*, *Science* **305**, 367 (2004).
2. Royal Society, *Ocean Acidification Due to Increasing Atmospheric Carbon Dioxide* (Policy Document 12/05, Royal Society, London, 2005).
3. V. J. Fabry, B. A. Seibel, R. A. Feely, J. C. Orr, *ICES J. Mar. Sci.* **65**, 414 (2008).
4. C. Langdon, M. J. Atkinson, *J. Geophys. Res.* **110**, C09S07 (2005).
5. U. Riebesell *et al.*, *Nature* **407**, 364 (2000).
6. L. Zondervan, *Deep-Sea Res. II* **54**, 521 (2007).
7. G. Langer *et al.*, *Geochem. Geophys. Geosyst.* **7**, Q09006 (2006).
8. M. D. Iglesias-Rodríguez *et al.*, *Science* **320**, 336 (2008).
9. E. Paasche, *Phycologia* **49**, 503 (2002).
10. D. M. Iglesias-Rodríguez, O. M. Schofield, J. Batley, L. K. Medlin, P. K. Hayes, *J. Phycol.* **42**, 526 (2006).
11. J. C. Cubillos *et al.*, *Mar. Ecol. Prog. Ser.* **348**, 47 (2007).
12. J. B. Ries, D. C. McCorkle, A. L. Cohen, *Eos Trans. AGU*, Ocean Sciences Meeting, OS003, abstract 2884 (2008).
13. J. C. Orr *et al.*, *Nature* **437**, 681 (2005).
14. T. Tyrrell, B. Schneider, A. Charalampopoulou, U. Riebesell, *Biogeosci. Discuss.* **4**, 3581 (2007).
15. R. A. Feely, C. L. Sabine, J. M. Hernandez-Ayon, D. Ianson, B. Hales, *Eos Trans. AGU*, Ocean Sciences Meeting, OS003, abstract 1185 (2008).

Published online 17 April 2008;
10.1126/science.1157130.
Include this information when citing this paper.

MATERIALS SCIENCE

Stronger, Tougher Steels

Insights into the microstructure and brittle failure of steels is leading to a new generation of structural steels.

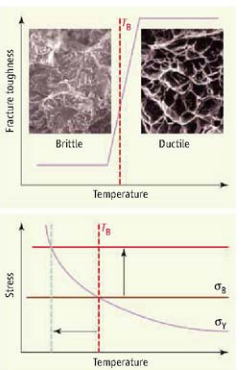
J. W. Morris Jr.

Steel is the workhorse of our infrastructure. Stronger, tougher steels are always needed to reduce weight and improve safety in transportation, enhance architectural flexibility in construction, and improve performance in heavy machinery. For structural steels to be both strong and tough (resistant to fracture), they must not be used at temperatures below the ductile-brittle transition temperature, T_B, at which the steel loses its toughness and fractures in a brittle mode (see the figure, top panel). This transition results from a competition between plastic deformation and brittle fracture at the tips of cracks or flaws in the steel. It can be controlled by techniques such as grain refinement that inhibit brittle fracture, or by techniques such as controlled delamination that facilitate plastic deformation. On page 1057 of this issue, Kimura *et al.* (*1*) show how these approaches can be combined to achieve low T_B and high toughness in an ultrahigh-strength low-alloy steel.

The ductile-brittle transition became infamous during the Second World War when hundreds of mass-produced American "liberty ships" cracked or split open in northern waters (*2*). This aquatic carnage led to the first systematic research on the ductile-brittle transition in steel and stimulated the development of fracture mechanics as a science.

After the war, the need for cold-weather and cryogenic structures—from pipelines, ships, and offshore structures for use in cold waters to the extremely low temperatures required for transporting and storing liquid helium—impelled research on methods to lower T_B. Initial successes used relatively expensive alloy additions, particularly nickel. As techniques such as transmission electron microscopy began to reveal the complex microstructures of high-strength steels, it became clear that the alloy additions modified the microstructure in ways that promoted ductile fracture (*3*). Since the 1970s, researchers have focused on thermal or thermomechanical processes that produce desirable microstructures in low-alloy steels, or processes that combine alloying and thermomechanical treatments to achieve high strength and toughness.

The source of the ductile-brittle transition can be understood from the Yoffee diagram (see the figure, bottom panel), an idealized representation of the relative likelihood of ductile or brittle fracture at the tip of a preexisting crack (*4*). As the applied stress is increased toward failure, the peak stress at the crack tip first reaches one of two levels: the effective yield stress (σ_Y) (*5*), at which significant plastic deformation occurs, or the brittle fracture stress (σ_B), at which the crack propagates in a brittle mode. The effective yield stress rises as the temperature drops, and the ductile-brittle transition temperature, T_B, is that at which σ_Y rises above σ_B.



Avoiding brittle fracture. (Top) At the ductile-brittle transition (T_B), the steel becomes brittle and most of its toughness is lost. **(Bottom)** The Yoffee diagram illustrates T_B at the intersection of the effective yield stress (σ_Y) and the brittle fracture stress (σ_B). Raising the brittle fracture stress (vertical arrow) causes a substantial decrease in T_B (horizontal arrow).

The Yoffee diagram suggests two ways to suppress the ductile-brittle transition: raising the brittle fracture stress or lowering the effective yield stress. Because a high yield stress is a desirable feature of a structural steel, most research has focused on the brittle fracture stress. Raising the brittle fracture stress (vertical arrow in the figure) is very effective in lowering T_B (horizontal arrow).

To raise the brittle fracture stress, it is first necessary to suppress intergranular fracture along grain boundaries. Intergranular embrittlement is usually a result of chemical impurities such as sulfur and phosphorus that accumulate between the grains. It can be controlled by purifying the steel to remove the offending species, or by introducing microalloy additions to "getter" them into mechanically innocuous precipitates. Suppressing intergranular fracture in steel lowers T_B but does not eliminate the ductile-brittle transition. The steel can also fracture by a cleavage mechanism in which the individual crystal grains break by separation on the face planes of the iron crystal; this transgranular cleavage mechanism must also be addressed.

The best way to raise σ_B for transgranular cleavage is to refine the "effective grain size," that is, to decrease the mean free path of a cleavage crack before a grain boundary or

Department of Materials Science and Engineering, University of California, Berkeley, CA 94720, USA. E-mail: jwmorris@berkeley.edu

other crystallographic discontinuity forces it to deviate or stop. Processes to do so include thermomechanical treatments during hot-rolling from slab to plate, sheet, or bar (*6*), thermal cycling treatments that manipulate the transformations that occur during quenching (*7*), and tempering treatments that interpose fine distributions of nanosized second-phase particles to disrupt the intragranular crystallography (*7*).

The complexity of the many possible microstructures of steel often makes it difficult to decipher what the effective grain size really is. Recently, electron backscatter diffraction (EBSD) has been used to map crystallographic patterns in the microstructure of steels and to identify the effective grain sizes in several important microstructures (*4*, *8*). This information is being systematically exploited in steel research, particularly in various "supersteel" projects in Asia (*9*).

The second generic approach to lowering T_B is to decrease the effective yield stress at the crack tip. A simple way to do this is to replace thick beams with a laminate of thin sheets, a technique used by 19th-century engineers in early steels. A modern alternative is to design the microstructure so that it spontaneously delaminates in the stress field at the crack tip, effectively dividing itself into a laminate of thin sheets and lowering the effective stress at the crack tip by a factor of 2 or more (*10*). This approach was used (but, to my knowledge, never published) by manufacturers of steel line pipe more than 20 years ago. It lowers T_B but also lowers the toughness in the ductile mode, presumably because the weaker of the delaminated sheets fracture under small loads. Particularly in the case of ultrahigh-strength steels, high toughness is required for structural reliability in service.

Kimura *et al.* now describe a clever method for combining grain refinement and delamination to obtain a particularly promising combination of properties. The steel is processed to high strength through a combination of grain refinement and precipitation hardening. The grain refinement is accomplished by bar-rolling, producing a fibrous texture that delaminates ahead of and perpendicular to the propagating crack tip, thereby relaxing the crack-tip stress concentration without producing planes of weakness to lower toughness in the ductile mode.

Defeating the ductile-brittle transition in steel is always difficult, and as the figure makes clear, this becomes particularly challenging for an ultrahigh-strength steel. As Kimura *et al.* show, this challenge can be met with low-alloy steel by tailoring the microstructure to achieve high strength while resisting brittle fracture. As advances in the understanding of brittle fracture and the control of microstructure feed into the design of new structural steels, more exciting results of this kind can be expected.

References and Notes

1. Y. Kimura, T. Inoue, F. Yin, K. Tsuzaki, *Science* **320**, 1057 (2008).
2. A. J. McEvily, *Metal Failures* (Wiley-Interscience, New York, 2001), chap. 1.
3. G. Thomas, *Metall. Trans. A* **2**, 2373 (1971).
4. J. W. Morris Jr., C. S. Lee, *Z. Gao, ISIJ Int.* **43**, 410 (2003).
5. The effective yield stress in the Yoffee diagram is the stress required to plastically deform the severely constrained material at the tip of a sharp crack. Unless the loaded piece is very thin, geometric constraints at the crack tip have the consequence that plastic deformation is difficult there, and requires a stress σ_y that is three to five times the conventional yield stress measured in standard tensile tests using unconstrained samples and tabulated in tables of engineering properties (*4*, *11*).
6. S. Takaki, T. Maki, Eds., *Ultrafine Grained Steels* (Iron and Steel Institute of Japan, Tokyo, 2001).
7. Z. Gao, C. S. Lee, J. W. Morris Jr., *Acta Mater.* **54**, 5511 (2006).
8. T. Maki, in *Fundamentals of Martensite and Bainite: Toward Future Steels with High Performance*, T. Furuhara, K. Tsuzaki, Eds. (Iron and Steel Institute of Japan, Tokyo, 2007), pp. 1–10.
9. *The Development of High Performance Steels for the 21st Century* (Pohang Iron and Steel Company Ltd., Pohang, Korea, 2002).
10. K. T. Venkateswara Rao, W. Yu, R. O. Ritchie, *Metall. Trans. A* **20A**, 485 (1989).
11. R. M McMeeking, G. M. Parks, in *Elastic-Plastic Fracture* (American Society for Testing and Materials, Philadelphia, 1979), pp. 175–194.

12. Supported by NSF grant DMR 0304629.

10.1126/science.1158994

MOLECULAR BIOLOGY

Slicing and Dicing for Small RNAs

James A. Birchler and Harsh H. Kavi

A new type of small RNA and mode of gene regulation is discovered in fly and mammals.

Nearly two decades after gene silencing was first described (*1*), the field continues to reveal new and diverse roles for the small RNA moieties involved. These functions include modulating the translation of messenger RNA (mRNA) into protein, establishing chromosomal architecture, regulating stem cell renewal, and providing defense against viruses and mobile genetic elements (transposons) that could cause deleterious mutations (*2*). Three major classes of small RNAs have been defined in plants and animals: microRNAs (miRNAs), small interfering RNAs (siRNAs), and Piwi-interacting RNAs (piRNAs). Previous work has suggested that dedicated pathways generate each class of small RNA. Now, six recent studies, including one by Ghildiyal *et al.* on page 1077 of this issue, reveal an interplay of these canonical pathways in generating a new class of endogenous siRNAs (endo-siRNA) (*3*–*8*) (see the figure).

miRNAs are encoded by the genome and processed to 22–base pair products that interact with homologous mRNAs to modulate their translation and affect developmental processes. siRNAs are 21 base pairs long and generated from double-stranded RNA precursors such as those from viruses and endogenous transposons. Recently, piRNAs have been described in the germ line of flies and mammals (*9*). They are approximately 25 base pairs long with homology to transposable elements and other sequences.

In the fly *Drosophila melanogaster*, the enzyme Dicer-1 cleaves double-stranded miRNA precursors, and the resulting products bind to the protein Argonaute1 (Ago1). siRNAs are generated by Dicer-2 from double-stranded RNA precursors, and bind to Argonaute2 (Ago2), whose "slicer" activity cleaves homologous RNAs. piRNAs are generated by a dicer-independent mechanism that relies on the slicer function of a separate clade of argonaute proteins consisting of Piwi, Aubergine, and Argonaute3 (*9*). These proteins use antisense transcripts (complementary to mRNA) encoded by piRNA clusters in the genome that harbor transposable element fragments, and target the destruction of transposon sense transcripts. New studies report sequencing of extensive pools of small RNAs from various somatic and germline sources in flies and mice (*3*, *4*, *6*–*8*) and identify a new type of siRNA that is homologous to transposons and protein-encoding gene sequences.

Ghildiyal *et al.* found that although most small RNAs in fly somatic cells are miRNAs, some have characteristics of siRNAs—equal quantities of sense and antisense orientations, a 21–base pair length, and 2′-*O*-methyl modifications at the 3′ end. Most

Division of Biological Sciences, University of Missouri, Columbia, MO 65211, USA. E-mail: birchlerj@missouri.edu

of these endo-siRNAs are homologous to transposable elements, and many originate from piRNA clusters. Curiously, mutations of the genes encoding Dicer-2 (*dcr-2*) and Ago2 (*ago2*) did not eliminate all of these small endo-siRNAs, suggesting that Dicer-1 and Ago1 can generate and bind, respectively, some siRNAs under these circumstances. Moreover, somatic pi-like RNAs, which exhibit characteristics of the germline namesake, were found in *ago2* mutants, raising the possibility of somatic activity of the Piwi family of argonautes. Mutations in *dcr-2* or *ago2* increased the expression of certain transposons, thus linking endo-siRNAs to the control of transposable elements in the soma.

Similarly, Czech *et al.* (*4*) found endo-siRNAs in somatic and germline cells of flies by analyzing sequences from three sources: those bound to Ago1 and Ago2, those present in small RNA libraries, and those from flies with mutations in the canonical miRNA and siRNA pathways. They observed that the protein Loquacious (Loqs), which associates with Dicer-1 for miRNA production, also interacts with Dicer-2 and is required for some endo-siRNA formation. Okamura *et al.* (*5*) analyzed the processing of products of endogenous long RNA hairpin (hp) genes in flies. These hpRNAs are processed by Dicer-2, in conjunction with Loqs, and then bind to Ago2. The siRNAs generated from these genes can silence selected mRNAs using the slicer function of Ago2 rather than by translational repression. Collectively, these new fly studies illustrate the interplay among mechanisms generating various classes of small RNAs.

In a complementary study, Kawamura *et al.* (*6*) sequenced small RNAs bound by Ago2 in cultured *Drosophila* cells. The endo-siRNAs recovered showed homology to a subset of transposons. Depletion of Dicer-2 and Ago2 in these cells reduced the quantities of endo-siRNAs and increased the expression of the corresponding transposons. Endo-siRNAs with homology to transposons were also found in fly ovaries, and their function depended on Dicer-2 and Ago2, revealing that the germ line has both siRNA- and piRNA-generating machineries.

The production of small RNAs by slicing and dicing has also been reported in mouse germline cells. Watanabe *et al.* (*7*) and Tam *et al.* (*8*) sequenced small RNAs from mouse oocytes and found both siRNAs and piRNAs. Depletion of Dicer and Ago2 in these cells reduced the number of siRNAs while increasing that of transposons and certain mRNAs. Because there is only one *Dicer* gene in mammals, these results demonstrate that this enzyme functions in both miRNA and siRNA production. The piRNAs found in oocytes were bound to MILI, a mammalian counterpart of Piwi. In this case, piRNAs and siRNAs sometimes correspond to distinct transposons, indicating preferred mechanisms for their generation, although transposon-rich loci could also give rise to both siRNAs and piRNAs.

SMALL RNAs

Germ line

Type	Length (nucleotides)	Mechanism of generation	Biological role
miRNA	21–23	Dicer dependent	Regulation of translation
piRNA	24–27	Dicer independent, Piwi clade dependent	Transposon regulation and unknown functions
endo-siRNA	21–23	Dicer dependent	Transposon regulation

Soma

miRNA	21–23	Dicer-1 dependent (in *Drosophila*)	Regulation of translation
piRNA like	24–27	In Ago2 mutants (in *Drosophila*)	Transposon regulation
endo-siRNA	21–23	Dicer-2/Ago2 dependent (in *Drosophila*)	Regulation of transposons, mRNAs, and heterochromatin
hpsiRNA	21–23	Dicer-2/Ago2/Loqs dependent (in *Drosophila*)	Regulation of mRNAs

Small RNAs. A summary of small RNAs found in germline and somatic cells and their identified functions.

It is not unexpected that endo-siRNAs to transposable elements would be found in animals, given their presence in plants (*10*) and fungi (*11*). Based on the sophisticated immune systems of vertebrates, early thinking posited that the small RNA–processing machinery involved Dicer primarily for the maturation of miRNAs. The discovery of a piRNA system that does not involve Dicer, but protects the germ line from transposon expression, reinforced this view. Although it is difficult to compare the strength of gene silencing across species, *Drosophila* somatic cells require multiple copies of a transgene to reduce the expression of homologous transcripts (*12*), and the effect is not as great as seen with a few copies of many transgenes in plants (*2*). Perhaps the evolution of robust immune systems in animals relaxed the selection pressure for a strong response of the siRNA defense against viruses. Thus, the piRNA system might represent a germline–accentuated mechanism to ameliorate the mobility of transposable elements, and thus reduce the occurrence of mutations. However, the finding that endo-siRNAs are present with homology to complementary pseudogenes and homologous genes as a newly recognized mode of gene regulation (*8*), and that long hpRNA genes produce siRNAs that can affect gene expression (*5*), might suggest that these roles for siRNAs have maintained the Dicer-dependent machinery at some level in both somatic and germline cells of animals. This machinery could still metabolize double-stranded viral and transposon RNAs.

Dual small RNA–processing mechanisms in animals might reflect a response to transposons that evolve to evade silencing by one or the other mechanism in the never-ending arms race between transposable elements and the host genome. Transposons in mouse oocytes have somewhat distinct profiles of siRNAs or piRNAs, indicating that certain elements can avoid one of the mechanisms of silencing. Any transposon that escapes both systems will likely proliferate into new genomic positions and produce substrates that could be subject to one or the other type of silencing. The piRNA loci are graveyards of transposable elements whose ghosts in the form of small RNAs commit fratricide against their homologous brethren. From a selfish DNA point of view, it is interesting to contemplate how the transposons are "tricked" by the host to insert into piRNA loci and in turn whether classes of transposons evolve mechanisms to avoid these locations for insertion.

References

1. M. A. Matzke *et al.*, *EMBO J.* **8**, 643 (1989).
2. P. J. Paddison, P. K. Vogt, Eds., *RNA Interference*, Current Topics in Microbiology and Immunology (Springer, Berlin, 2008), vol. 320.
3. M. Ghildiyal *et al.*, *Science* **320**, 1077 (2008); published online 10 April 2008 (10.1126/science.1157396).
4. B. Czech *et al.*, *Nature* 10.1038/nature07007 (2008).
5. K. Okamura *et al.*, *Nature* 10.1038/nature07015 (2008).
6. Y. Kawamura *et al.*, *Nature* 10.1038/nature06938 (2008).
7. T. Watanabe *et al.*, *Nature* 10.1038/nature06908 (2008).
8. O. H. Tam *et al.*, *Nature* 10.1038/nature06904 (2008).
9. A. A. Aravin, G. J. Hannon, J. Brennecke, *Science* **318**, 761 (2007).
10. A. Hamilton, O. Voinnet, L. Chappell, D. Baulcombe, *EMBO J.* **21**, 4671 (2002).
11. A. Chicas, C. Cogoni, G. Macino, *Nucleic Acids Res.* **32**, 4237 (2004).
12. M. Pal Bahdra, U. Bhadra, J. A. Birchler, *Mol. Cell* **9**, 315 (2002).

Published online 8 May 2008;
10.1126/science.1159018
Include this information when citing this paper.

RETROSPECTIVE

Edward N. Lorenz (1917–2008)

A pioneer of chaos theory had a remarkable ability to distill complex systems to their physical essence.

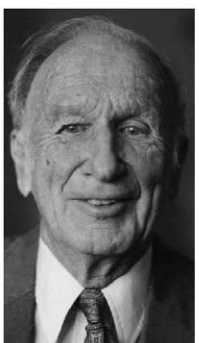
Kerry Emanuel

To students taking their first physics course, the idea that the evolution of the material universe is governed by physical laws may come as a revelation. Hitherto mysterious events are unmasked as predictable, and the same laws that govern the rise and set of the Sun and Moon can be used to place a probe into a precise region of Jupiter's atmosphere, over 600 million km from Earth. One is led quickly to the clockwork universe of René Descartes, in which every event, every outcome, is predetermined by events at the beginning of the universe.

Cracks in this deterministic world view began to show up early in the 20th century. In 1906, Henri Poincaré, after the frustrating experience of failing to solve the three-body orbital problem, conjectured that the evolution of some systems is so sensitive to initial-state perturbations that prediction becomes practically impossible. Twenty years later, Werner Heisenberg and the Copenhagen School showed that the subatomic world can only be described in probabilistic terms. Yet the physics of macroscopic phenomena continued to proceed under the mantra of strict determinism, encouraged by—among other developments—the stunning success of general relativity. It was left to a mild-mannered MIT mathematician to hammer the last nail into the casket of the Cartesian universe.

Edward Norton Lorenz was born in West Hartford, Connecticut, on 23 May 1917, and developed an early interest in science. "As a boy I was always interested in doing things with numbers and was also fascinated by changes in the weather," he later wrote (*1*). After receiving degrees in mathematics from Dartmouth in 1938, and Harvard in 1940, he enlisted in the U.S. Army Air Corps, where he served as a weather forecaster. After the war, he studied meteorology at MIT, earning a doctorate in 1948 and then joining the faculty of the meteorology department there.

The story of how Lorenz embarked on the development of the modern theory of chaos is by now well told. Working with a primitive computer in 1961, he was calculating numerical solutions to a set of three differential equations representing a reduction of the equations governing fluid convection in a box. Wanting to carry the integration further in time, he restarted a calculation at about the midpoint of his first run, using the numerical output as his starting state. Escaping the racket of the machine, he stepped out for a cup of coffee, but on returning found that the solution had diverged greatly from the first run. At first suspecting a machine malfunction, he quickly realized that he had stumbled on a proof of Poincaré's conjecture: On reentering the data, he had merely rounded the output to three significant figures.



In 1963, Lorenz reported and interpreted these results in a paper entitled "Deterministic nonperiodic flow" in the *Journal of the Atmospheric Sciences*. The article went almost unnoticed outside the atmospheric sciences for nearly a decade. This and subsequent work on the mathematical properties of chaotic systems has been called the third scientific revolution of the 20th century.

Chaos is not merely about the practical problem of prediction. Lorenz and others showed that there are mathematical (and almost certainly physical) systems for which the predictability horizon—the time limit over which a prediction can be made—approaches a finite limit as the initial error approaches zero. Such systems are formally unpredictable beyond this finite horizon, no matter how precise the computation and the specification of the initial state. Even the motion of a frictionless billiard ball becomes completely unpredictable after only 11 collisions, owing to the uncertainty principle's limit on describing its initial state.

In the field of atmospheric science, Lorenz is as well known for his work on the general circulation and dynamics of the atmosphere as he is for chaos theory. His 1955 paper developing the concept of available potential energy in fluids is widely regarded as a classic, as is his 1969 treatise on the general circulation of the atmosphere. By the early 1960s, Lorenz had succeeded in explaining the sudden transitions in the flow of fluid in "rotating dishpan" experiments, which are crude laboratory analogs of the general circulation of the atmosphere. This represented a major advance in understanding, accomplished with a remarkably simplified set of equations. The mathematical distillation of complex physical systems to their physical essence, so evident in this work, was the hallmark of Lorenz's technique.

Lorenz was a member of the MIT faculty until his retirement in 1987. He continued research in mathematics and meteorology thereafter, proofing a scientific paper only days before his death on 16 April 2008.

Ed Lorenz was an exceptionally unassuming and kind man who loved nothing better than hiking and cross-country skiing with his friends, colleagues, and family, to whom he was completely devoted. He was a great teacher, much beloved by his students, and won the MIT Meteorology Department's graduate teaching award every year it was offered. His teaching emphasized the value of rigor, clarity, simplicity, and logic, all very much in evidence in his scientific writing.

Although a very practical man, he was not indifferent to the philosophical implications of his work. In his masterful final book, *The Essence of Chaos*, he writes: "We must wholeheartedly believe in free will. If free will is a reality, we shall have made the correct choice. If it is not, we shall still not have made an incorrect choice, because we shall not have made any choice at all, not having a free will to do so" (*2*). Its logical construction and careful choice of wording make it a fitting remark by which to remember Lorenz.

Program in Atmospheres, Oceans, and Climate, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139, USA. E-mail: emanuel@texmex.mit.edu

CREDIT: MIT NEWS OFFICE

References and Notes

1. From a biographical sketch date-stamped 4 October 1996, housed at MIT News Office.
2. E. N. Lorenz, *The Essence of Chaos* (Univ. of Washington Press, WA, 1993), p. 160.

10.1126/science.1159438

INTRODUCTION

Lost in Microbial Space

LONG AGO, WE REALIZED THAT THE NUMBER OF STARS IN THE FIRMAMENT WAS uncountable, and the light they shed is from the far distant past. Our planet's microbial universe is equally awesome: Thinking about the number of organisms living in one human gut is discombobulating, and we are probably inhaling spores that have been blowing in the wind for years. Advances in genome-sequencing technology now allow previously uncultured organisms to be probed in their natural environments, whether they live intimately with us in our guts or in the far, deep sludge of the ocean's abyss (see Roussel *et al.*, p. 1046). This special issue examines some of the challenges in understanding the significance of all this diversity to the patterns and dynamics of arguably the most important organisms on Earth.

Although the organisms remain unseen, Falkowski *et al.* (p. 1034) point out that the microbial world drives some of the largest-scale phenomena on the planet, including photosynthesis, nitrogen cycling, and pandemics of infectious disease. In the opening Editorial, Tiedje and Donohue (p. 985) comment on ways in which this catalytic power might be harnessed to supply the needs of human societies.

Rapid-sequencing technology has also revealed unimaginable degrees of diversity. Paradoxically, the more diversity we find among microbes, the less simple it becomes to assign microorganisms to species. In a News story (p. 1031), Bohannon examines the phylogenetic entanglements caused by dizzying exchanges among the genomes of bacteria, archaea, and their viruses and discusses ongoing efforts to classify microbes with respect to their environments. Certainly, the fascinating microcosms harbored by marine sponges seem to defy biogeography. In a News story, Vogel (p. 1028) describes the remarkable similarity of microbial communities in sponges throughout the world. Nevertheless, Green *et al.* (p. 1039) argue that rather than investigate the distribution of individuals, we should instead be looking for the geographic distribution of traits.

Despite its uniformity to the human eye, the open sea offers not just a chance for microorganisms to swim but also niches generated by changes in light levels and salinity, marine "snot" and more resistant particles, and homes on the surfaces of animals. Hunt *et al.* (p. 1081) investigate the effect of niche partitioning in open water in a report showing that the habitat preferences of marine vibrios have taxon-specific patterns. But there are yet more layers of complexity in open-sea communities, and Strom (p. 1043) reviews the interdependencies of microbes in the plankton. Potentially, interrelationships can reach deep scales of detail; for example, Andersson and Banfield (p. 1047) have traced the path of genetic debris left by viruses preying on microbes in biofilms. Other examples of mechanisms used to forge microbial relationships are explored in *Science Signaling* (www.sciencemag.org/microbialecology/): Dow discusses how closely related plant pathogens use similar signaling molecules in distinct ways, Lee describes how microbes in the gut dampen the immune response, and Wartha and Henriques-Normark discuss how mast cells make extracellular traps to kill microbes.

The shock of finding seemingly endless variation in microbial genomes has thus quietly revolutionized thinking about evolution and the species concept, as well as strengthened realization of the importance of the ecological services supplied by the microbial universe.

– CAROLINE ASH, JOHN FOLEY, ELIZABETH PENNISI

Microbial Ecology

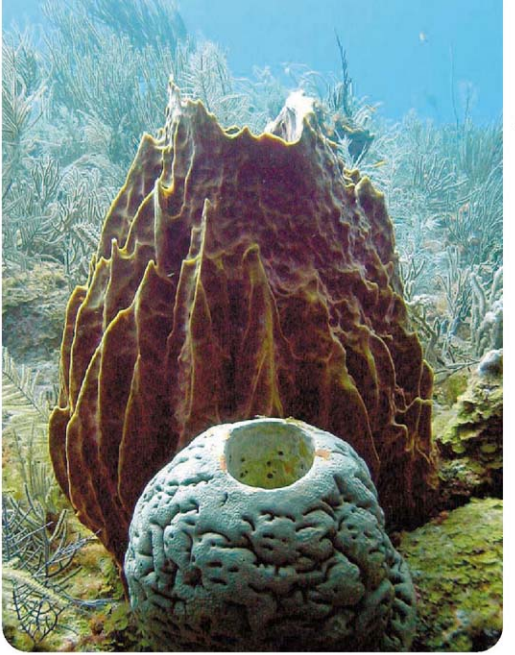
CONTENTS

News

Reviews

See also related Editorial page 985; News of the Week story page 1001 and News Focus story page 1006; Science Express Report by R. E. Ley et al.; Brevia page 1046; Research Article page 1047; Report page 1081; and Science Signaling material on page 979 or at www.sciencemag.org/microbialecology/.

Science



Hospitable habitat. The brown tube sponge *Agelas conifera* (foreground) and the giant barrel sponge *Xestospongia muta* both have complex microbial communities living among their cells.

CREDIT: JOSEPH PAWLIK/UNIVERSITY OF NORTH CAROLINA AT WILMINGTON

NEWS

The Inner Lives of Sponges

Symbiotic ties, bioactive compounds, and mysterious distributions of bacteria characterize these ancient invertebrates

A SPONGEFUL OF BACTERIA IS THE LAST thing a dishwasher wants to think about. But for Jörn Piel, the more microbes he finds in a sponge, the better. Not a synthetic one, of course, but those that adorn tropical reefs and populate the ocean bottom.

One of evolution's more ancient animals, sponges at first glance seem quite simple—little more than loose consortiums of semi-autonomous cells, stuck in one place filtering food from the water column. But a closer look reveals a surprising twist. "With many species, under the microscope you see almost exclusively bacteria" among the cells, says Piel, an organic chemist at the University of Bonn in Germany. Just as microbial ecologists are demonstrating the extent and importance of microbes in ecosystems as diverse as guts and glaciers (see p. 1046), Piel and others are slowly uncovering a hidden microbial world inside sponges.

It's a difficult job, as almost none of the sponges' inhabitants grow in the lab. But through genetic studies, researchers are revealing the rich diversity and unusual distribution of these microbes. Some investigators are pinning down the roles bacteria play in sponge biology and ecology. The microbes are teaching scientists about evolution, symbiosis, and the mind-boggling variety of life on our planet. "It never ceases to amaze me that a sponge, an organism that just sits there and pumps bucketfuls of water through its canals," has such a rich and varied, yet highly specific, inner life, says marine microbiologist Michael Taylor of the University of Auckland in New Zealand. The research also has a practical side: Piel and others are betting that sponge-dwelling bacteria could be the source of potentially valuable compounds for treating cancer, malaria, and other human diseases.

Sharper focus

The first hint of the sponges' pervasive inhabitants came in the 1960s and '70s, as new equipment allowed longer and deeper dives that gave researchers their first up-close look at the diversity of life on the ocean bottom. It quickly became clear that something else was living among the sponges' cells. Looking at the first electron microscope images of sponge tissue, marine biologist Jean Vacelet and his colleagues at the University of Marseille spotted what looked like a half-dozen different types of bacteria. Other researchers "took sponges and squeezed them out over culture plates to see what would grow," recalls marine ecologist Robert Thacker of the University of Alabama, Birmingham, but it was difficult to follow up on the finds. At most, 5% of sponge-dwelling species have thrived in the lab, says microbial ecologist Ute Hentschel of the University of Würzburg in Germany. And the sponges themselves "are incredibly hard to keep alive," Thacker says.

Therefore, Hentschel, Thacker, and others have been using indirect methods to piece together a picture of this reclusive community. Most of the evidence comes from studies of the gene for 16*S* ribosomal RNA (rRNA), a piece of the genome that scientists use to identify unknown microbes in the environment. Differences in this gene can serve as a useful measure of the kinship between two species.

These genetic studies uncovered a distinctive and extensive community, identifying more than 100 species of microbes that are found in sponges but not in the surrounding water. This distribution indicates that these bugs are long-term residents rather than passersby. An individual sponge might host dozens of different species, and overall, the molecular analyses have found an impressive variety: 14 bacterial phyla, two phyla of archaea, and several types of eukaryotic microbes.

Such diversity initially suggested multiple, independent acquisitions of microbial symbionts. But evidence is building that sponges of different types and in different oceans host strikingly similar microbial communities. Hentschel and her colleagues showed in 2002 that sponges from the coast of Japan, the Red Sea, the Mediterranean, and the Republic of Palau in the South Pacific contained microbes that are more closely related to each other than to the microbes in the seawater from which the sponges were harvested. "It's astounding," says Susanne Schmitt, a postdoc in Hentschel's lab. The different sponges the scientists sampled diverged millions of years ago, she says, but they are home to very similar, and very complex, microbial communities. In 2007, Taylor and his colleagues found the same result when they analyzed the entire database of 16*S* rRNA sequences available from sponge-dwelling microbes collected from all over the world—nearly 2000 sequences in all.

But Taylor, Hentschel, and their colleagues are still trying to work out what the results mean. Microbes might have colonized a sponge early in the group's evolutionary history and acquired characteristics that enabled them to live in sponges full-time, Taylor proposes. Those sponge-loving microbes could have then spread to other sponges—and other oceans. And such a scenario could explain what may be a new phylum called *Poribacteria*, after Porifera, Latin for "sponge." *Poribacteria* have been found throughout the world, albeit exclusively in sponges.

Fruitful partnership

As with much of microbial ecology, the sponge specialists have been focused primarily on taking a census. "I go in just trying to figure out what's there—what people did collecting insects in the forest 100 years ago," Taylor explains. But he and his colleagues are now starting to take the next step, because census data can't tell researchers what each side gets out of the relationship. Ecologists want to know if the microbes and their hosts are obligate symbionts, unable to survive without each other, or whether the microbes are tolerated but dispensable guests, says Michael Wagner, a microbial ecologist at the University of Vienna in Austria: "If we want to understand these communities, we have to know the function each member plays."

Yet even after decades of study, scientists are still not exactly sure what sponges and their microbes are doing for each other. Living in nutrient-poor but sunlit waters in the lagoons of the Republic of Palau, sponges of the family Dysideidae are home to blue-green algae that probably provide their hosts with energy and carbon. The sheer mass of the microbes may help support the meter-high giant barrel sponge *Xestospongia muta*, in which bacteria can sometimes make up 40% of a sponge's volume. Microorganisms may even help defend their hosts against disease-causing bacteria.

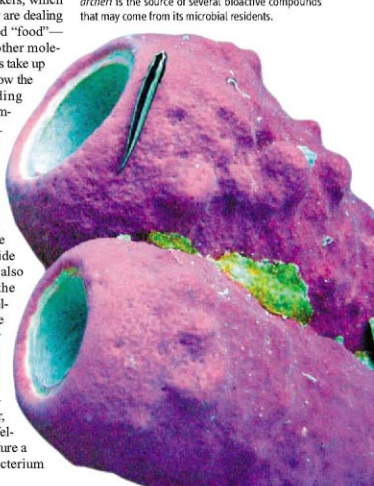
But those are educated guesses rather than proven observations. "And there are a whole lot of other things that are going on that we just don't know about," says molecular ecologist Russell Hill of the University of Maryland Biotechnology Institute in Baltimore.

To try to get a picture of the daily goings-on inside a sponge, Wagner and his colleagues are catching sponge microbes in the act of "eating." The researchers have just started experiments on several species of sponges that host *Poribacteria*. No *Poribacteria* have ever been cultured in the lab, but the scientists are able to keep the host sponges alive in aquaria, at least for a short time. They use a technique that allows them to observe the metabolic activity of individual microbes and sponge cells. They expose the sponge to fluorescently labeled rRNA markers, which lets them know what species they are dealing with, and to radioactively labeled "food"—amino acids, bicarbonate, and other molecules. They then watch which cells take up the labeled morsels and follow how the morsels are processed, including whether the sponge consumes compounds excreted by the microbes. "We're asking not only 'Who are you?' but also 'What are you eating?'" he says.

Whatever their function, the microbes seem important enough for sponges to pass on to future generations. In the female sponge, nurse cells, which provide the "yolk" for developing eggs, also ferry blue-green algae from the sponge's outer layers to the developing oocytes located deeper in the sponge matrix. In 2005, Kayley Usher and her colleagues at the University of Western Australia in Perth even found blue-green algae in the sperm of the sponge *Chondrilla australiensis*. A year later, Julie Enticknap, a postdoctoral fellow in Hill's lab, was able to culture a sponge-dwelling alphaproteobacterium from the larvae of a sponge collected off the coast of Florida, another indication of possible parent-to-offspring transmission.

But that study highlights what may be the most baffling mystery in sponge microbiology. Usually when symbionts are passed from parent to offspring, the partners undergo what is called cospeciation, and the microbes develop a unique genetic signature and become confined to that particular host. "But that doesn't happen here," says Hentschel. The bacteria in the larvae proved closely related to those cultured from unrelated sponges growing in Jamaica, Indonesia, and the Chesapeake Bay in the United States. The best explanation for the broad distribution of this bacterium—and for many other species found across the globe—she says, is that sponges acquire their resident bacteria both from their parents and from the environment.

To date, no sponge-specific microbe has turned up in seawater, but scientists have a distinct disadvantage when it comes to sampling. Although a 1-kilogram sponge can filter thousands of liters of seawater a day, Hills says that "if we are lucky, we



Medicinal potential. The stovepipe sponge *Aplysina archeri* is the source of several bioactive compounds that may come from its microbial residents.

CREDIT: JOSEPH PAWLIK/UNIVERSITY OF NORTH CAROLINA AT WILMINGTON

filter 200 liters," so the chances of finding an uncommon microbe, such as the larvae's alphaproteobacterium, are small.

If sponges are taking microbes in from the surrounding environment, they need to be able to tell friend from foe from food. And the microbes need a way to protect themselves against accidental or intentional rebuff by their hosts. Electron microscope images reveal that most sponge-dwelling bacteria have either thickened cell walls or slime capsules that might prevent the sponge cells from digesting them. Once established, these resident microbes, or the sponge itself, seem to produce chemicals that discourage interlopers. Several antibiotic compounds isolated from sponges efficiently kill bacteria found in the water column but do not affect sponge-dwelling organisms.

Sea-based drugs

Those antibiotic compounds are driving at least some of the interest in sponges. For bioprospectors looking for potential new drugs from the sea, "sponges are one of the best sources of bioactive compounds," says Hill. The chemical from which the antiretroviral drug AZT was derived was first found in a Caribbean sponge. In the lab, other compounds from sponges kill cancer cells and malaria parasites.

AIDS, cancer, and malaria are not the sponge's concern, but a powerful chemical defense arsenal is, points out microbial ecologist Julie Olson of the University of Alabama, Tuscaloosa: "Sponges can't evade predators, and if something blocks their [water-filtering] passages, it's a death sentence." To protect against unfriendly microbes, a successful sponge probably needs a range of chemical weapons, she says.

Initially, marine biochemists using the "grind and find" approach assumed that most of those chemicals came from the sponge itself. But as the diversity of the sponges' residents became clear, many began to suspect that at least some of the compounds might come from the lodgers rather than the hosts. The quest to come up with enough of a bioactive compound for clinical testing is proving that these suspicions are well-founded.

To date, few sea-based drugs have made it to the clinic. "Supply is the primary reason there is no blockbuster so far," says Piel. It's been almost impossible to purify a compound in quantities large enough for animal and human testing, and the chemical structures are often too complex for large-scale chemical synthesis. Halichondrin B, for example, is a powerful antitumor compound in lab tests. But scientists calculated that clinical trials would require at least 10 grams of the substance. The best producer, a New Zealand sponge group called *Lissodendoryx*, yielded 300 milligrams per metric ton of sponges. Because the entire population of *Lissodendoryx* was estimated at 280 metric tons, it was clear that harvesting from the wild was not sustainable, Piel says.

Piel is trying to get around that problem. His lab is fishing for genes involved in making promising compounds. The goal is to find the set of genes that codes for the potential drug's synthesis, and if the original host won't grow in the lab, to transfer those genes to a microbe that is happy in an artificial environment. The designer microbe would then pump out enough of the drug for testing.

The technique is pointing to bacteria as the source of many of the compounds. "So far, every time we've found a gene cluster, we found typical bacterial genes" nearby, Piel says. In 2004, the group reported that they had pinned down the genes responsible for producing compounds called polyketides in a dark-red sponge called *Theonella swinhoei* that lives in coral reefs. (In the lab, polyketides can kill tumor cells, and several types are in clinical trials.)

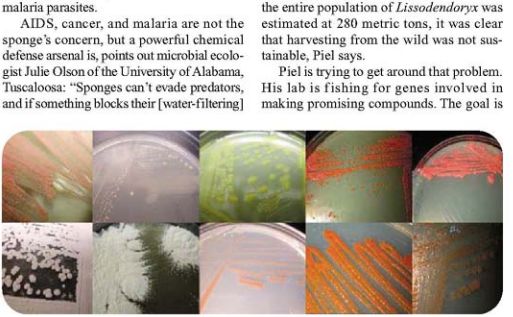
Based on the genes' similarity to known genes, Piel and his colleagues concluded that the genes most likely come from a still-uncultured microbe. What's odd, however, is that these genes are quite similar to polyketide genes belonging to a bacterium that lives in the guts of beetles. The researchers are currently working to transfer the sponge microbe's genes to a lab-friendly host.

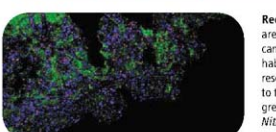
Hill and his group have focused on trying to harness the original bacteria producers. "Part of the problem is that people have in their heads that all symbionts are difficult to grow," Hill says. But patience and hard work can pay off. "Sometimes we get new colonies after months of incubation."

In recent work, Hill's lab has homed in on the source of a particularly promising compound called manzamine A. In lab tests, manzamines kill malaria parasites more efficiently than either chloroquine or artemisinin, two of the leading antimalarial drugs. The compound was first identified in a sponge collected off the coast of Okinawa, but related compounds have since turned up in dozens of unrelated sponge species all over the world—a strong hint, Hill says, that they are produced by a microbe shared by all these species.

In as-yet-unpublished work, his group has isolated the bacterium that produces manzamine A. The microbe should give scientists their first steady supply of the compound, allowing them to make and test new derivatives, Hill says. Such studies led to the eventual development of AZT, Hill points out. And he is hoping sponges—or at least their microbes—will again lend a hand in the fight against deadly disease.

–GRETCHEN VOGEL





Reclusive residents. These Actinobacteria (*above*) are some of the few sponge-dwelling microbes that can grow in culture. To find out more about the habits of bacteria that don't grow in the lab, researchers use tagged genes and labeled nutrients to trace their fates. Here (*left*), sponge cells are green, nitrite-oxidizing symbionts of the genus *Nitrospira* are red, and all other bacterial symbionts appear blue.

CREDITS: TOP SERIES: MATTHEW ANDERSON/UNIVERSITY OF MARYLAND BIOTECHNOLOGY INSTITUTE; (FLUORESCENT IMAGE) KRISTIN BERGAUER, JULIAN STOECKER, AND MICHAEL WAGNER/UNIVERSITY OF VIENNA

Sex in the salt pond. The search for genetically isolated microbial special in hypersaline pools like this one near San Diego, California, revealed rampant gene swapping among species.

NEWS

Confusing Kinships

Understanding microbial evolution and ecology rests on a solid classification system, but coming up with one is difficult

ALONG THE PARCHED SLOPES OF A CANYON in Israel, in salty pools in Spain and Algeria, and on countless petri plates in laboratories around the world, a scientific debate is playing out: how to classify microbial organisms accurately. Not only are bacteria and their ilk amazingly diverse, but genes cross species lines so frequently that researchers argue whether microbial species exist at all. At stake is much more than the esoteric record-keeping of taxonomists, says R. Thane Papke, a microbiologist at the University of Connecticut, Storrs: "This is about our fundamental understanding of evolution." Without rigorously categorizing diversity, "we're really stuck."

The problem is that prokaryotes, the single-celled organisms without a nucleus, are promiscuous. Instead of one cell splitting into two genetically identical daughter cells, over and over, most take part in a global orgy of gene swapping, passing genes between different taxa. This spells trouble for traditional systematics, built as it is on the assumption that organisms' genes faithfully reveal their common ancestry. Whereas most genes in a particular microbe do come from its direct ancestor, many may not, making lines of descent difficult, if not impossible, to describe (*Science*, 1 May 1998, p. 672).

This is not the first time that microbial systematics has been disrupted. The tremors began 4 decades ago when DNA sequence became the gold standard for classifying organisms, revolutionizing our understanding of how microbes fit into life's grand scheme. Now, the availability of whole genomes and DNA sequence from complete microbial communities is shaking up the field anew. "An earthquake is coming for microbial systematics," says Hans-Peter Klenk, a microbiologist at the German Collection of Microorganisms and Cell Cultures (DSMZ) in Braunschweig. And after the dust settles, what microbial "species" will look like is anyone's guess.

What's in a name?

Classifying microbes has never been easy. Well into the 20th century, bacteria were considered members of the fungi, themselves erroneously classified as plants. At the first International Microbiological Congress in Paris in 1930, scientists decided that microbes needed their own scheme. At that time, members of the new Commission on Nomenclature and Taxonomy called microorganisms "in part plants, in part animals, and in part primitive." They therefore concluded that these single-celled creatures belonged with neither.

From then on, microbiologists assigned names to the organisms they found wriggling and dividing under their microscopes based on the few characteristics that could be reliably observed. Whether microbes irreversibly soak up a dye called crystal violet designated them as Gram-negative—such as our common gut inhabitant, *Bacteroides fragilis*—or Gram-positive, such as the food-poisoning microbe *Clostridium botulinum*. The *Bacillus* species were distinguished by their need for oxygen and ability to form spores. Others, such as *Streptococcus pneumoniae*, earned their monikers based on the diseases they caused. "But everyone understood that these species definitions were rather arbitrary," says Klenk, because there was no way to confirm that they reflected genetic relatedness. Nonetheless, "it was a practical system that could help microbiologists know what they were talking about."

And help it did. Microbiology exploded during the second half of the 20th century, transforming every field it touched, including medicine, agriculture, ecology, and even geology. The enterprise was built on an ever-growing microbial family tree and "type cultures" of each microbial species, representative batches kept in laboratories around the world. Type cultures made it possible for researchers to replicate and build on previous experiments, says Klenk.

By the late 1970s, there were some 40,000 type cultures. "And that's when we had our first big shock," says Klenk. Using the newly available tools of molecular biology, scientists compared DNA sequence from different microbial species and found that "we were completely wrong about evolutionary relationships." The prokaryotes split, some staying in the familiar Bacteria and others shifting into Archaea, single-celled organisms that are superficially similar to

CREDIT: THANE PAPKE/UNIVERSITY OF CONNECTICUT, STORRS

bacteria but whose genetic architecture is more like our own.

Starting in 1980, the naming of microbial species went in for a complete overhaul by the International Committee on Systematics of Prokaryotes. The standards for type cultures were made far more rigorous, says Klenk, calling for DNA sequence data and more thoroughly documented isolation. The 40,000 type cultures were pared down to about 4000—growing since then to 6800.

But DNA studies have continued to muddy the microbial waters. "In some ways, the more genetic data arrives, the less clear things get," says Christophe Fraser, an epidemiologist at Imperial College London.

One problem revealed by DNA is the vastness of microbial diversity, says Fred Cohan, a microbiologist at Wesleyan University in Middletown, Connecticut. "All of microbiology was based on what we could culture in the lab," with heavy emphasis on pathogens. But lab-culturable bugs turn out to be "certainly less than 1%" of living microbes. As technologies improved, ever more DNA sequence has been harvested from environmental samples, representing all the microbial genetic material in a pinch of soil or milliliter of seawater. The realization that "a soil community contains tens of millions of bacteria" is "humbling," says Cohan.

Sequencing of whole genomes has presented microbiologists with an even more daunting challenge: Microbes have an active "sex" life. Scientists have long known that genes can move between microbes; the spread of antibiotic-resistance genes since the 1950s is a case in point. "But what's surprising is how frequent and widespread it turns out to be," says Fraser.

For traditional species to be well-defined, their genes need to flow vertically, from parents to offspring and nowhere else. But among microbes, genes can move along a bewildering variety of routes between genomes: sliding through bridges between cellular membranes, hitchhiking inside viruses, or even getting sucked up from the environment as naked fragments. That means that any given microbe can have a large number of "parents" from many different species. But "if the genes are moving freely, then how can you nonarbitrarily define the relations between different microbes?" says Papke. "Do you really have species at all?"

The earthquake begins

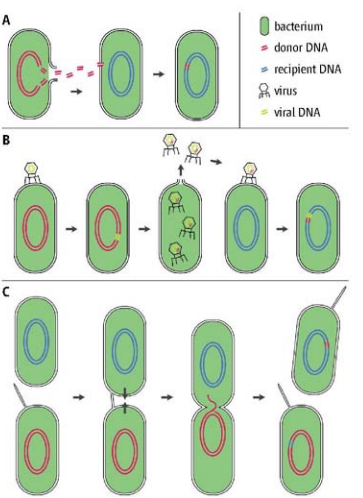
Yet Fraser and others haven't given up on classifying microbes. They think there are other ways to define "species."

Take recombination, for example. In plants and animals that have no choice but to reproduce sexually, recombination happens every generation: Matching strands of DNA on chromosomes line up and swap segments, producing offspring with a shuffled deck of genes from each parent. This can happen within microbes, too. When foreign DNA finds its way inside a microbe's cell membrane, it has a chance of lining up with a similar sequence and swapping segments with the genome. Such recombination events happen infrequently, but when they deliver a more useful version of a gene, the recombined genome can even sweep through a population to become the norm. DNA sequencing of many microbes from the same population has revealed that recombination is much more common than was ever thought.

But the "crucial insight" is that the frequency of recombination depends on the kinship between the donor and recipient microbes, says Fraser. The more closely related two microbes are—and hence, the greater the similarity between their genomes—the greater the chance that recombination will happen. When the microbes are too different, "the recombination rate drops off steeply," he says, effectively blocking gene flow. "If we can define this threshold, then that could be a rigorous way to define microbial species."

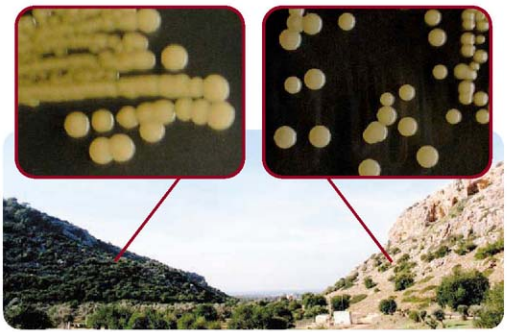
Last year, Fraser and colleagues used computer simulations to study microbial evolution (*Science*, 26 January 2007, p. 476). The results suggested that when the rate of recombination is high enough, genetically isolated groups can emerge that are "analogous" to traditional species, he says.

Recent evidence for Fraser's view has come from a study of real-world microbial recombination (*Science*, 11 April, p. 237). In this case, a species is being lost. A team led by Martin Maiden, a microbial geneticist at the University of Oxford, U.K., studied two species of *Campylobacter* bacteria. The genomes of *C. coli* and *C. jejuni* share only 86.5% of their most conserved DNA sequence, due to millions of years of adaptation to different wild host animals. But in the 10,000 years since the advent of agriculture, the two species have been living together cheek-by-jowl in farm animals, and there, recombination is on the rise, blurring what in other environments is a clear species line. Because recombination is occurring nearly 20 times faster in the genome of *C. coli*, that genome is becoming ever more



DNA highway. Genes have many options for moving between different microbes, including getting slurped up as fragments from the environment (A), hitchhiking inside retroviruses (B), and getting swapped with similar sequence on a foreign genome after a cellular tryst called conjugation (C).

CREDIT: J. NEWFIELD/*SCIENCE*



Defined by lifestyle. In this Israeli canyon, researchers have divided one bacterial "species" into "ecotypes." *Bacillus simplex* ecospecies Graminifolius (*left*) prefers the grassy southern slope, whereas *B. simplex* ecospecies Sylvaticus prefers the dry and sunny northern slope.

CREDITS (TOP TO BOTTOM): (COLONIES) JOHANNES SIKORSKI/DSMZ; (TREE) CONAN LIFELEVAN UNIVERSITY

like *C. jejuni*'s, and Maiden predicts that the two will eventually become indistinguishable.

But other tests of recombination as a gold standard for microbial species have not yielded such clear results. Papke was part of a team that studied microbes in three hypersaline pools, two in Spain and a third 250 kilometers away in Algeria. The salt-loving microbes isolated in each pool, all members of the genus *Halorubrum*, should represent separate species. Yet among the 153 strains tested, there was a bewildering degree of recombination—even between "species" in pools separated by the Mediterranean, Papke and his colleagues reported last August in the *Proceedings of the National Academy of Sciences* (*PNAS*). Considering their blush-worthy promiscuity, "we'll just have to accept that microbes evolve in ways that don't allow them to be pigeonholed into species," says Papke.

The eco-challenge

But some researchers argue that similarity in lifestyle, not just genes, is the way to classify microbes. Even with DNA flowing willy-nilly, microbes pigeonhole themselves into coherent groups by adapting to different niches, says Cohan: "The key to understanding microbial diversity is ecology." Cohan and others would like to do away with microbial "species" as the "fundamental unit" of diversity. Instead, microbes would be divided into "ecotypes," based first on genetic relatedness and more finely on shared adaptations to a particular habitat.

To demonstrate the existence of ecotypes among real-world microbes, Cohan and a team led by David Ward, a microbial ecologist at Montana State University in Bozeman, have studied *Bacillus* bacteria from a group of arid canyons in Israel. The bacteria have adapted to the canyons' various microenvironments, says Cohan, from the harsh, dry northern slope to the relatively mild, lush southern slope, and a periodically flooded streambed between. The team isolated 218 bacteria from these locations, all members of the genus *Bacillus*, and sequenced the DNA of five genes from each. On the basis of their highly conserved 16*S* ribosomal RNA gene sequences, bacteria from one canyon divided neatly into two species, *B. simplex* and *B. licheniformis*. But stopping there "would not be informative," says Cohan, because the sequences of other genes revealed substantial diversity between bacteria with identical 16*S* sequence. That diversity is driven by adaptation to different microhabitats, says Cohan.

So Ward and Cohan's team devised a method called "ecotype simulation" to categorize the canyon microbes based on lifestyle. First, they generated a phylogenetic tree based on variation in the four marker genes that were different in the various bacteria. Then, the simulation comes up with putative ecotypes: It divides the bacteria into clusters of genetically similar individuals, all the while checking to make sure that the similarities are due to adaptive changes and not to chance. Finally, the researchers use what is known about the actual environment of each microbe—solar exposure, moisture, soil type, plant resources—to test whether the predicted ecotypes correspond to various microhabitats.

The team identified as many as 30 distinct ecotypes across all canyons studied. Rather than species, these are the "fundamental units" of microbial diversity, the team concluded in a paper published in *PNAS* 19 February. Cohan says they are now "preparing to propose" some of these ecotypes, such as *Bacillus simplex* ecospecies Graminifolius, to the systematics community for formal recognition.

Ward and Cohan aren't the only researchers using ecology to make sense of microbial diversity. A study of marine microbes led by Martin Polz at the Massachusetts Institute of Technology in Cambridge found that groups of bacteria of the same species occupy different parts of the plankton community, and even during specific seasons (see p. 1081).

But an ecology-based classification faces an uphill battle for acceptance. "I just don't see how ecotypes can work," says Papke. He thinks gene swapping is so frequent among microbes that recombination, not ecological adaptation, is the main cause of diversity. Ecotypes "still have a way to go," agrees Ford Doolittle, a microbiologist at Dalhousie University in Halifax, Canada.

Doolittle is pessimistic that anything better than a "compromise solution" can ever be achieved for microbial systematics. When conditions are just right, microbes "may cluster into what we could all agree to call species," says Doolittle, on the basis of either "ecotypes" or rates of recombination. "But there is no reason to suppose that conditions will often or even ever be right and thus no reason to suppose that there must be 'fundamental units' of bacterial diversity." Traditional systematics requires such units, he says, "but needing something to be true does not make it so."

And how will all this affect day-to-day microbiology? "I expect species names and type cultures will continue," says Klenk. Papke agrees: "We need to be able to have a conversation." But as for the "fundamental unit" of microbial diversity and what, if anything, is represented by microbial species, "we'll probably need philosophers to sort that out."

—JOHN BOHANNON

REVIEW

The Microbial Engines That Drive Earth's Biogeochemical Cycles

Paul G. Falkowski,[1]* Tom Fenchel,[2]* Edward F. Delong[3]*

Virtually all nonequilibrium electron transfers on Earth are driven by a set of nanobiological machines composed largely of multimeric protein complexes associated with a small number of prosthetic groups. These machines evolved exclusively in microbes early in our planet's history yet, despite their antiquity, are highly conserved. Hence, although there is enormous genetic diversity in nature, there remains a relatively stable set of core genes coding for the major redox reactions essential for life and biogeochemical cycles. These genes created and coevolved with biogeochemical cycles and were passed from microbe to microbe primarily by horizontal gene transfer. A major challenge in the coming decades is to understand how these machines evolved, how they work, and the processes that control their activity on both molecular and planetary scales.

Earth is ~4.5 billion years old, and during the first half of its evolutionary history, a set of metabolic processes that evolved exclusively in microbes would come to alter the chemical speciation of virtually all elements on the planetary surface. Consequently, our current environment reflects the historically integrated outcomes of microbial experimentation on a tectonically active planet endowed with a thin film of liquid water (*1*). The outcome of these experiments has allowed life to persist even though the planet has been subjected to extraordinary environmental changes, from bolide impacts and global glaciations to massive volcanic outgassing (*2*). Although such perturbations led to major extinctions of plants and animals (*3*), to the best of our knowledge, the core biological machines responsible for planetary biogeochemical cycles have survived intact.

The explosion of microbial genome sequence data and increasingly detailed analyses of the structures of key machines (*4*) has yielded insight into how microbes became the biogeochemical engineers of life on Earth. Nevertheless, a grand challenge in science is to decipher how the ensemble of the core microbially derived machines evolved and how they interact, and the mechanisms regulating their operation and maintenance of elemental cycling on Earth. Here we consider the core set of genes responsible for fluxes of key elements on Earth in the context of a global metabolic pathway.

Essential Geophysical Processes for Life

On Earth, tectonics and atmospheric photochemical processes continuously supply substrates and remove products, thereby creating geochemical cycles (*5*, *6*). These two geophysical processes allow elements and molecules to interact with each other, and chemical bonds to form and break in a cyclical manner. Indeed, unless the creation of bonds forms a cycle, planetary chemistry ultimately will come to thermodynamic equilibrium, which would lead inevitably to a slow depletion of substrates essential for life on the planetary surface. Most of the H_2 in Earth's mantle escaped to space early in Earth's history (*7*); consequently, the overwhelming majority of the abiotic geochemical reactions are based on acid/base chemistry, i.e., transfers of protons without electrons. The chemistry of life, however, is based on redox reactions, i.e., successive transfers of electrons and protons from a relatively limited set of chemical elements (*6*).

The Major Biogeochemical Fluxes Mediated by Life

Six major elements—H, C, N, O, S, and P—constitute the major building blocks for all biological macromolecules (*8*). The biological fluxes of the first five of these elements are driven largely by microbially catalyzed, thermodynamically constrained redox reactions (Fig. 1). These involve two coupled half-cells, leading to a linked system of elemental cycles (*5*). On geological time scales, resupply of C, S, and P is dependent on tectonics, especially volcanism and rock weathering (Fig. 1). Thus, biogeochemical cycles have evolved on a planetary scale to form a set of nested abiotically driven acid-base and biologically driven redox reactions that set lower limits on external energy required to sustain the cycles. These reactions fundamentally altered the surface redox state of the planet. Feedbacks between the evolution of microbial metabolic and geochemical processes create the average redox condition of the oceans and atmosphere. Hence, Earth's redox state is an emergent property of microbial life on a planetary scale. The biological oxidation of Earth is driven by photosynthesis, which is the only known energy transduction process that is not directly dependent on preformed bond energy (*9*).

The fluxes of electrons and protons can be combined with the six major elements to construct a global metabolic map for Earth (Fig. 2). The genes encoding the machinery responsible for the redox chemistry of half-cells form the basis of the major energy-transducing metabolic pathways. The contemporary pathways invariably require multimeric protein complexes (i.e., the microbial "machines") that are often highly conserved at the level of primary or secondary structure. These complexes did not evolve instantaneously, yet the order of their appearance in metabolism and analysis of their evolutionary origins are obscured by lateral gene transfer and extensive selection. These processes make reconstruction of how electron transfer reactions came to be catalyzed extremely challenging (*10*).

In many cases, identical or near-identical pathways may be used for the forward and reverse reactions required to maintain cycles. For example, methane is formed by methanogenic Archaea from the reduction of CO_2 with H_2. If the hydrogen tension is sufficiently low, however, then the reverse process becomes thermodynamically favorable; methane is oxidized anaerobically by Archaea closely related to known, extant methanogens that apparently use co-opted methanogenic machinery in reverse. Low hydrogen tension occurs when there is close spatial association with hydrogen-consuming sulfate reducers (*11*–*13*); thus, this process requires the synergistic cooperation of multispecies assemblages, a phenomenon that is typical for most biogeochemical transformations. Similarly, the citric acid cycle oxidizes acetate stepwise into CO_2 with a net energy yield. In green sulfur bacteria, and in some Archaebacteria, the same cycle is used to assimilate CO_2 into organic matter with net energy expenditure. Indeed, this may have been the original function of that cycle (*14*). Typically, in one direction, the pathway is oxidative, dissimilatory, and produces adenosine 5′-triphosphate, and in the opposite direction, the pathway is reductive, assimilatory, and energy consuming.

However, reversible metabolic pathways in biogeochemical cycles are not necessarily directly related, and sometimes are catalyzed by diverse, multispecies microbial interactions. The various oxidation and reduction reactions that drive Earth's nitrogen cycle (which, before humans, was virtually entirely controlled by microbes) are a good example. N_2 is a highly inert gas, with an atmospheric residence time of ~1 billion years. The only biological process that makes N_2 accessible for the synthesis of proteins and nucleic acids is nitrogen fixation, a reductive process that transforms N_2 to NH_4^+. This biologically irreversible reaction is catalyzed by an extremely conserved heterodimeric enzyme complex, nitrogenase, which is inhibited by oxygen (*15*). In the presence of oxygen, NH_4^+ can be oxidized to nitrate in a two-stage pathway, initially requiring a specific group of Bacteria or Archaea that oxidize ammonia to NO_2^- (via hydroxyamine), which is subsequently oxidized to NO_3^- by a different suite of nitrifying bacteria (*16*). All of the nitrifiers use the small differences in redox potential in the oxidation reactions to reduce CO_2 to

[1]Environmental Biophysics and Molecular Ecology Program, Institute of Marine and Coastal Sciences and Department of Earth and Planetary Sciences, Rutgers University, New Brunswick, NJ 08901, USA. [2]Marine Biological Laboratory, University of Copenhagen, Strandpromenaden 5, DK-3000 Helsingør, Denmark. [3]Department of Civil and Environmental Engineering and Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

*To whom correspondence should be addressed. E-mail: falko@marine.rutgers.edu (P.G.F); tfenchel@bio.ku.dk (T.F.); delong@mit.edu (E.F.D.)

organic matter (i.e., they are chemoautotrophs). Finally, in the absence of oxygen, a third set of opportunistic microbes uses NO_2^- and NO_3^- as electron acceptors in the anaerobic oxidation of organic matter. This respiratory pathway ultimately forms N_2, thereby closing the N cycle. Hence, this cycle of coupled oxidation/reduction reactions, driven by different microbes that are often spatially or temporally separated, forms an interdependent electron pool that is influenced by photosynthetic production of oxygen and the availability of organic matter (*17*).

Are the niches for all possible redox reactions occupied by microbial metabolism? Although some metabolic transformations, and the microbes that enable them, have been predicted to exist solely on the basis of thermodynamics, and only later were shown to actually occur (*18*, *19*), not all predicted pathways have been found. Some, such as the oxidation of N_2 to NO_3^-, may be too kinetically constrained for biological systems. Similarly, no known photosynthetic organism can photochemically oxidize NH_4^+.

Coevolution of the Metabolic Machines

Due to physiological and biochemical convenience, elemental cycles generally have been studied in isolation; however, the cycles have coevolved and influence the outcomes of each other. Metabolic pathways evolved to utilize available substrates produced as end products of other types of microbial metabolism, either by modification of existing metabolic pathways or by using established ones in reverse (*20*). Photosynthesis is another example of the evolution of multiple metabolic pathways that lead to a cycle. Typically, reduction and oxidation reactions are segregated in different organisms. In photosynthesis, the energy of light oxidizes an electron donor, i.e., H_2O in oxygenic photosynthesis and HS^-, H_2, or Fe^{2+} in anoxygenic

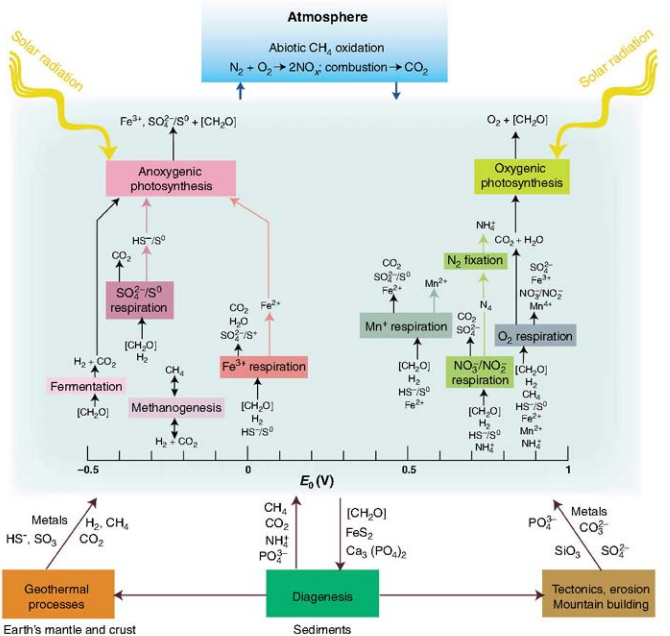


Fig. 1. A generalized biosphere model showing the basic inputs and outputs of energy and materials. Geochemical (abiotic) transformations are represented at the top (atmospheric) and bottom (tectonic and geothermal) compartments, while microbially driven biochemical processes are represented in the middle, biospheric compartment (in blue) and the sediments. Biological element cycling is not completely closed due to losses through sedimentation of organic carbon and nitrogen, carbonate, metal sulfides, sulfate, and phosphate, and losses to the atmosphere via denitrification. Regeneration of available forms of these elements is contingent on geological processes: erosion and geothermal activity. Electron acceptors (oxidants) in the respiratory processes have been arranged from left to right according to increasing capacity to accept electrons. The redox couples (at pH 7) for the reactions are approximate; the exact values depend upon how the individual reactions are coupled.

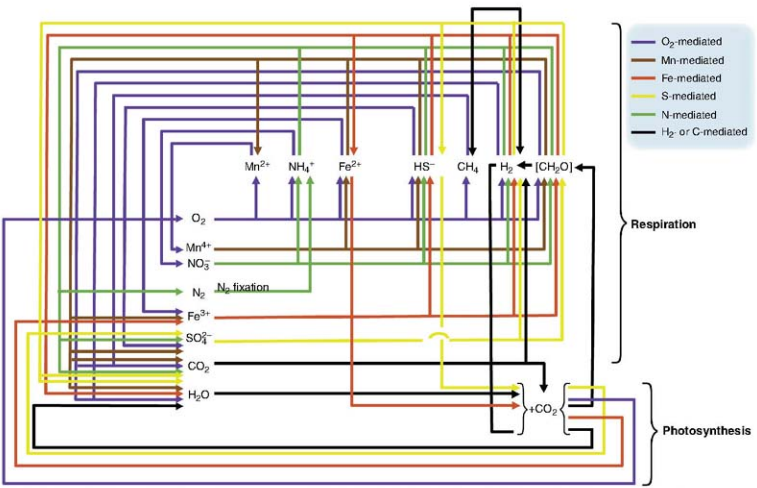


Fig. 2. A schematic diagram depicting a global, interconnected network of the biologically mediated cycles for hydrogen, carbon, nitrogen, oxygen, sulfur, and iron. A large portion of these microbially mediated processes are associated only with anaerobic habitats.

photosynthesis, and the electrons and protons generated in the process are used to reduce inorganic carbon to organic matter with the formation of higher-energy bonds. The resulting oxidized metabolites may in turn serve as electron acceptors in aerobic or anaerobic respiration for the photosynthetic organisms themselves or by other, nonphototrophic organisms that use these "waste products" as oxidants (*21*).

The outcome of the coupled metabolic pathways is that on geological time scales, the biosphere can rapidly approach relatively self-sustaining element cycling on time scales of centuries to millennia. On longer time scales, perpetuation of life remains contingent on geological processes and the constant flux of solar energy. Essential elements or compounds, such as phosphate, carbon (either as carbonate or organic matter), and metals, are continuously buried in sediments and are returned to the biosphere only through mountain building and subsequent erosion or geothermal activity (Fig. 1).

There is little understanding of how long it took for reaction cycles to develop from local events to global alteration of prevailing geologically produced redox set points. The last common ancestor of extant life presumably possessed genes for the adenosine triphosphatase complex required to maintain ion gradients generated by photochemical or respiratory processes. Regardless, one of the last metabolic pathways to emerge was oxygenic photosynthesis.

Oxygenic photosynthesis is the most complex energy transduction process in nature: More than 100 genes are involved in making several macromolecular complexes (*22*). Nevertheless, indirect evidence shows that this series of reactions had evolved by ~3 billion years ago (*23*), although the atmosphere and the upper ocean maintained a very low concentration of O_2 for the next ~0.5 billion years (*24*, *25*). The production and respiration of nitrate must have evolved after the advent of oxygenic photosynthesis, as there can be no nitrate without oxygen (*16*). Although the succession of probable events that led to the global production of O_2 is becoming increasingly clear (*26*, *27*), the evolutionary details delimiting important events for other redox cycles and elements are more ambiguous.

Attempts to reconstruct the evolution of major dissimilatory metabolic pathways are mainly based on geological evidence for the availability of potential electron donors and oxidants during the early Precambrian (*23*). Although we can gain some idea of the relative quantitative importance of different types of energy metabolism, we do not know the order in which they evolved. Indeed, the origin of life and the first reactions in energy metabolism probably never will be known with certainty. These events took place before any geogical evidence of life, and while phylogenetic trees and structural analyses provide clues regarding key motifs, so far they have not provided a blueprint for how life began. Stable-isotope fractionation has provided evidence for sulfate reduction and methanogenesis in 3.5-billion-year-old deposits (*28*), but these metabolic processes are presumably older.

Modes of Evolution

Molecular evidence, based on gene order and the distribution of metabolic processes, strongly suggests that early cellular evolution was probably communal, with promiscuous horizontal gene flow probably representing the principal mode of evolution (*29*). The distribution of genes responsible for the major extant catabolic and anabolic processes may have been distributed across a common global gene pool, before cellular differentiation and vertical genetic transmission evolved as we know it today. In the microbial world, not only individual genes but also entire metabolic pathways central to specific biogeochemical cycles appear to be frequently horizontally transferred; a contemporary analog is the rapid acquisition of antibiotic resistance in pathogenic bacteria (*30*). The dissimilatory sulfite reductases found in contemporary sulfate-reducing δ-proteobacteria, Gram-positive

bacteria, and Archaea are examples of horizontal gene transfer that reflect the lateral propagation of sulfate respiration among different microbial groups and environments (*31*). Indeed, with the exception of chlorophyll- or bacteriochlorophyll-based photosynthesis, which is restricted to Bacteria, and methanogenesis, which is restricted to representatives within the Archaea (*32*), individual bacterial and Archaeal lineages contain most major metabolic pathways. Even some of the molecular components of methanogens seem to have been laterally transferred to methane-oxidizing members of the domain Bacteria (*33*). Nitrogenases appear to have been transferred to oxygenic photosynthetic cyanobacteria late in their evolutionary history, probably from an Archaean source (*34*), and are widespread among diverse groups of Bacteria and Archaea (*35*). Ammonia monooxygenase genes that encode the key enzyme required for the oxidation of ammonia to hydroxylamine, a key step of the nitrogen cycle, are also widely distributed (*36*, *37*). Evidence also exists for lateral exchange of large "superoperons" encoding the entire anoxygenic photosynthetic apparatus (*38*). Presumably, severe nutritional or bioenergetic selective pressures serve as major drivers for the retention of horizontally transferred genes, thereby facilitating the radiation of diverse biogeochemical reactions among different organisms and environmental contexts.

Sequence Space Available

Although the absolute number of genes and protein families currently in existence is unknown, several approaches have been used to evaluate the relative depth of protein "sequence space" currently sampled. Microbial community genome sequencing (i.e., metagenomics) provides a cultivation-independent, and hence potentially less biased, view of extant sequence space. The number of protein families within individual Bacterial and Archaeal genomes depends linearly on the number of genes per genome, and hence genome size (*39*). The higher levels of gene duplication found in nonmicrobial eukaryotic genomes potentially allows them to escape this constraint and has resulted in different evolutionary strategies and genome organization (*39*). Regardless, genome size appears to be correlated with evolutionary rate, but not with core metabolic processes (*40*). So, what does the apparent diversity in microbial genomes signify?

Genome Diversity in Nature

To date, the rate of discovery of unique protein families has been proportional to the sampling effort, with the number of new protein families increasing approximately linearly with the number of new genomes sequenced (*41*). The size of protein families (the number of nonredundant proteins found within a family) among fully sequenced genomes follows a power law, with the greatest number of protein families containing only a few members (*39*). These trends among fully sequenced genomes are also mirrored in large-scale metagenomic shotgun sequencing efforts (*42*). Among the ~6 million newly predicted protein sequences from a recent ocean metagenomic survey, a total of 1700 new protein families were discovered with no homologs in established sequence databases. Even though this study increased the known number of protein sequences nearly threefold from just one specific habitat, the discovery rate for new protein families was still linear (Fig. 3). These data indicate that we have only just begun the journey of cataloguing extant protein sequence space.

The virtual explosion of genomic information has led to the hypothesis that there is limitless evolutionary diversity in nature. The vast majority of unexplored sequence space appears to encompass two categories of genes: a large and dynamic set of nonessential genes and pseudogenes, under neutral or slightly negative selective pressure (which we call "carry-on genes"), and a set of positively selected environment-specific gene suites, tuned to very particular habitats and organisms (which we call "boutique genes"). In contrast, the evolution of most of the essential multimeric microbial machines (including the basic energy transduction processes, nitrogen metabolic processes, ribosomes, nucleic acid replication enzymes, and other multienzyme complexes) is highly constrained by intra- and intermucleic acid, RNA-protein, protein-protein, protein-lipid, and protein–prosthetic group interactions (*22*), to the extent that even when the machines function suboptimally, they are retained with very few changes. For example, the D1 protein in the reaction center of Photosystem II, a core protein in the water-splitting reaction center found in all oxygenic photosynthetic organisms, is derived from an anaerobic purple bacterial homolog. During oxygenic photosynthesis, this protein is degraded by photooxidative cleavage approximately every 30 min (*43*). Rather than reengineer the reaction center to develop a more robust protein in the machine, a complicated repair cycle has evolved that removes and replaces the protein. Consequently, photosynthetic efficiency, especially at high irradiance levels, is not as high as theoretically possible (*44*), yet the D1 is one of the most conserved proteins in oxygenic photosynthesis (*22*). Similarly, nitrogenase is irreversibly inhibited by molecular oxygen, yet this core machine is also very highly conserved even though many nitrogen-fixing organisms live in an aerobic environment. To compensate, nitrogen-fixing organisms have had to develop mechanisms for protecting this enzyme from oxygen by spatially or tempo-

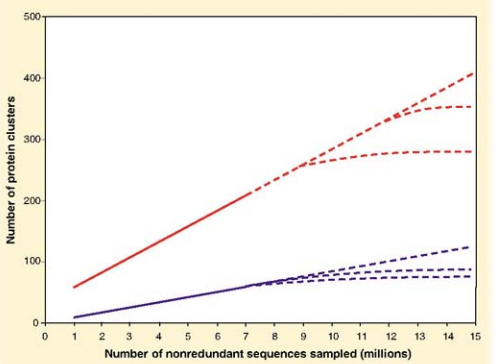


Fig. 3. Observed increases in new protein clusters with increasing sequence sampling [modified from Yooseph *et al.* (*42*)]. The number of new protein clusters discovered increases linearly with the number of nonredundant sequences sampled. We project hypothetical saturation profiles for the protein families. However, discovery of new protein families is much lower in protein clusters with greater membership. Seven different data sets of various sizes, including curated public databases and new data described in Yooseph *et al.* (*42*), were used to generate seven differently sized, nonredundant sequence data samples depicted. The red line shows protein clusters with ≥3 core sets of highly related sequences in a given cluster. The blue line shows protein clusters with ≥10 core sets of highly related sequences in a given cluster.

rally segregating nitrogen fixation from aerobic environments (*45*–*47*). In the contemporary ocean, ~30% of nitrogenase is nonfunctional at any moment in time, forcing overproduction of the protein complex to facilitate nitrogen fixation.

Is Everything Everywhere?

Abundant evidence exists for the rapid and efficient dispersal of viral particles and microbial cells, and for the genes they carry. At the same time, both microbial isolations and environmental genomic surveys indicate environmentally specific, quantitative distributional patterns of iron oxidation, methane metabolism, and photosynthesis (*11*, *48*, *49*). These distributions generally, but not always, reflect the environmental distributions of specific taxonomic groups. For example, the simplicity and modularity of rhodopsin-based photosynthesis appear to have led to the dispersal of this pathway into widely disparate taxonomic groups. The environmental distribution of these photoproteins therefore appears more reflective of habitat selective pressure than of any specific organismal or taxonomic distribution (*50*). Although the distributions of specific taxa may not vary greatly along a particular environmental gradient, in the absence of the relevant selection pressure, environmentally irrelevant genes may be lost rapidly (*51*).

The generalization that particular kinds of microbes always occur whenever their habitat requirements are realized is far from new (*52*). Although not necessarily metabolically active, viable bacteria of a particular functional type can be recovered from almost any environment, using appropriate types of enrichment cultures anywhere, even where that environment cannot support their growth. Hence, thermophilic bacteria can be grown from cold seawater (*53*), strict anaerobes from aerobic habitats (*54*), and microbial cells have been observed to accumulate in high numbers in surface snow at the South Pole (*55*). These observations can be explained by the sheer number of microbial cells occurring on Earth and consequent high efficiencies of dispersal and low probabilities of local extinction. Evidence for this also appears to be reflected in the vast number of very rare sequences revealed in rarefaction curves of deep microbial sampling surveys (*56*), which perhaps represents a sort of "biological detritus" from the very efficient microbial dispersal, coupled with extremely slow decay kinetics of individual microbial cells or spores in various resting states.

Very early in life's history the atmosphere and oceans were anoxic and the distribution of the first aerobic respiring microbes was confined to the close vicinity of cyanobacteria. By contrast, in the extant surface biosphere, aerobic conditions are very widespread. During the late Proterozoic (between ~750 and 570 million years ago) glaciations, large parts of Earth's surface may have been covered by ice, but even small remaining habitat patches will have assured the persistence and eventually dissemination of all types of prokaryotes. By extension, it is unlikely that mass-extinction episodes in the Phanerozoic (the past 545 million years), which strongly influenced the evolution of animals and plants, did not fundamentally influence the core metabolic machines. How then has the ancient core planetary metabolic gene set been maintained over the vast span of evolutionary time?

Microbes as Guardians of Metabolism

Dispersal of the core planetary gene set, whether by vertical or horizontal gene transfer, has allowed a wide variety of organisms to simultaneously, but temporarily, become guardians of metabolism. In that role, environmental selection on the microbial phenotype leads to evolution of the boutique genes that ultimately protect the metabolic pathway. If the pathway in a specific operational taxonomic unit does not survive an environmental perturbation, the unit will go extinct, but the metabolic pathway has a strong chance of survival in other units. Hence, the same selective pressures enabling retention of fundamental redox processes have persisted throughout Earth's history, sometimes globally, and at other times only in refugia, but able to emerge and exert ubiquitous selection pressure on ancillary genes. In essence, microbes can be viewed as vessels that ferry metabolic machines through strong environmental perturbations into vast stretches of relatively mundane geological landscapes. The individual taxonomic units evolve and go extinct, yet the core machines survive surprisingly unperturbed.

Humans may not yet be able to mimic the individual redox reactions that drive planetary processes; nevertheless, the interconnections between biogeochemical processes and the evolution of biologically catalyzed reactions are becoming more tractable for measurement and modeling. It is likely that the individual reactions that make life possible on Earth will be reasonably well described within the next few decades. Delineating how these machines coevolved and operate together to create the electron flows that predominate today on Earth's surface remains a grand challenge. Understanding biogeochemical coevolution is critical to the survival of humans as we continue to influence the fluxes of matter and energy on a global scale. Microbial life can easily live without us; we, however, cannot survive without the global catalysis and environmental transformations it provides.

References and Notes

1. V. Vernadsky, *Geochemistry and the Biosphere* (Synergetic Press, Santa Fe, 2007).
2. A. H. Knoll, *Life on a Young Planet: The First Three Billion Years of Evolution on Earth* (Princeton Univ. Press, Princeton NJ, 2003).
3. A. Hallam, P. Wignall, *Mass Extinctions and Their Aftermath* (Oxford Univ. Press, Oxford, 1997).
4. www.pdb.org/
5. A. Kluyver, H. Donker, *Chem. Zelle Gewebe* **13**, 134 (1926).
6. R. Williams, J. Frausto da Silva, *The Natural Selection of the Chemical Elements* (Clarendon, Oxford, 1996).
7. D. C. Catling, M. W. Claire, *Earth Planet. Sci. Lett.* **237**, 1 (2005).
8. W. H. Schlesinger, *Biogeochemistry: An Analysis of Global Change* (Academic Press, New York, ed. 2, 1997).
9. P. G. Falkowski, L. Godfrey, *Proc. R. Soc. London B Biol. Sci.*, 10.1098/rstb.2008.0054 (2008).
10. E. Allmer-White, M. J. Russell, *Biol. Direct* **3**, (2008).
11. S. J. Hallam *et al.*, *Science* **305**, 1457 (2004).
12. V. J. Orphan, C. H. House, K. U. Hinrichs, K. D. McKeegan, E. F. DeLong, *Science* **293**, 484 (2001).
13. B. Schink, *Antonie Van Leeuwenhoek* **81**, 257 (2002).
14. G. Wachtershauser, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 200 (1990).
15. J. R. Postgate, *Nitrogen Fixation* (Cambridge Univ. Press, Cambridge, ed. 3, 1998).
16. P. Falkowski, *Nature* **387**, 272 (1997).
17. K. Fennel, M. Follows, P. G. Falkowski, *Am. J. Sci.* **305**, 526 (2005).
18. E. Broda, *Z. Allg. Mikrobiol.* **17**, 491 (1977).
19. M. Jetten *et al.*, *FEMS Microbiol. Rev.* **22**, 421 (1998).
20. J. A. Gerlt, P. C. Babbitt, *Annu. Rev. Biochem.* **70**, 209 (2001).
21. R. E. Blankenship, M. T. Madigan, C. E. Bauer, Eds., *Anoxygenic Photosynthetic Bacteria* (Kluwer Scientific, Dordrecht, Netherlands, 1995).
22. T. Shi, T. S. Bibby, L. Jiang, A. J. Irwin, P. G. Falkowski, *Mol. Biol. Evol.* **22**, 2179 (2005).
23. D. E. Canfield, *Annu. Rev. Earth Planet. Sci.* **33**, 1 (2005).
24. J. Farquhar, H. M. Bao, M. Thiemens, *Science* **289**, 756 (2000).
25. A. Bekker *et al.*, *Nature* **427**, 117 (2004).
26. R. Blankenship, S. Sadekar, J. Raymond, in *Evolution of Primary Producers in the Sea*, P. G. Falkowski, A. H. Knoll, Eds. (Academic Press, San Diego, 2007), pp. 21–35.
27. J. Raymond, *Rev. Mineral. Geochem.* **59**, 211 (2005).
28. D. Canfield, R. Raiswell, *Am. J. Sci.* **299**, 697 (1999).
29. C. R. Woese, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 8742 (2002).
30. J. P. Gogarten, W. F. Doolittle, J. G. Lawrence, *Mol. Biol. Evol.* **19**, 2226 (2002).
31. M. Klein *et al.*, *J. Bacteriol.* **183**, 6028 (2001).
32. C. R. Woese, *Microbiol. Rev.* **51**, 221 (1987).
33. L. Chistoserdova, J. A. Vorholt, R. K. Thauer, M. E. Lidstrom, *Science* **281**, 99 (1998).
34. T. Shi, P. Falkowski, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 2510 (2008).
35. K. J. Kechris, J. C. Lin, P. J. Bickel, A. N. Glazer, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 9584 (2006).
36. S. J. Hallam *et al.*, *PLoS Biol.* **4**, e95 (2006).
37. M. Konneke *et al.*, *Nature* **437**, 543 (2005).
38. K. V. P. Nagashima, A. Hiraishi, K. Shimada, K. Matsuura, *J. Mol. Evol.* **45**, 131 (1997).
39. A. J. Enright, V. Kunin, C. A. Ouzounis, *Nucleic Acids Res.* **31**, 4632 (2003).
40. M. Oliver, D. Petrov, D. Ackerley, P. G. Falkowski, O. Schofield J. *Genome Res.* (2007).
41. V. Kunin, I. Cases, A. J. Enright, V. de Lorenzo, C. A. Ouzounis, *Genome Biol.* **4**, 401 (2003).
42. S. Yooseph *et al.*, *PLoS Biol.* **5**, e16 (2007).
43. I. Obud, D. J. Kyle, C. J. Arntzen, *J. Cell Biol.* **99**, 481 (1984).
44. S. P. Long, S. Humphries, P. G. Falkowski, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **45**, 633 (1994).
45. P. W. Ludden, G. P. Roberts, in *Anaygenic Photosynthetic Bacteria*, R. E. Blankenship, M. T. Madigan, C. E. Bauer, Eds. (Kluwer, Dordrecht, Netherlands, 1995), pp. 929–947.
46. I. Berman-Frank *et al.*, *Science* **294**, 1534 (2001).
47. R. C. Burns, R. W. F. Hardy, *Nitrogen Fixation in Bacteria and Higher Plants* (Springer, New York, 1975).
48. E. F. DeLong *et al.*, *Science* **311**, 496 (2006).
49. G. W. Tyson *et al.*, *Nature* **428**, 37 (2004).
50. N. U. Frigaard, A. Martinez, T. J. Mincer, E. F. DeLong, *Nature* **439**, 847 (2006).
51. E. Lerat, V. Daubin, H. Ochman, N. A. Moran, *PLoS Biol.* **3**, e130 (2005).
52. L. G. M. Baas Becking, *Geobiologie of Inleiding tot de Milieukunde* (Van Stockum and Zoon, The Hague, Netherlands, 1934).
53. M. F. Isaksen, F. Bak, B. B. Jørgensen, *FEMS Microbiol. Ecol.* **14**, 1 (1994).
54. M. Bianchi, J.-L. Teyssié, B. B. Fowler, *Mar. Ecol. Prog. Ser.* **88**, 55 (1992).

55. E. J. Carpenter, S. Lin, D. G. Capone, *Appl. Environ. Microbiol.* **66**, 4514 (2000).
56. M. L. Sogin *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 12115 (2006).

57. Supported by NASA and the Agouron Foundation (P.G.F.); the Danish Natural Science Research Council and The Carlsberg Foundation (T.F.); and the Gordon and Betty Moore Foundation, the Agouron Foundation, NSF, and U.S. Department of Energy (E.F.D.).

10.1126/science.1153213

REVIEW

Microbial Biogeography: From Taxonomy to Traits

Jessica L. Green,[1]* Brendan J. M. Bohannan,[1] Rachel J. Whitaker[2]

The biogeographic variation of life has predominantly been studied using taxonomy, but this focus is changing. There is a resurging interest in understanding patterns in the distribution not only of taxa but also of the traits those taxa possess. Patterns of trait variation shed light on fundamental questions in biology, including why organisms live where they do and how they will respond to environmental change. Technological advances such as environmental genomics place microbial ecology in a unique position to move trait-based biogeography forward. We anticipate that as trait-based biogeography continues to evolve, micro- and macroorganisms will be studied in concert, establishing a science that is informed by and relevant to all domains of life.

And so it was indeed: she was now only ten inches high, and her face brightened up at the thought that she was now the right size for going though the little door into that lovely garden.
Lewis Carroll (1865)

Imagine Carl Linnaeus in Alice's shoes, shrinking to only 10 micrometers high. Afforded the opportunity to investigate biological diversity at this spatial scale, would Linnaeus have remained committed to plant exploration, or would he have turned his attention to microbial life? It is not surprising that Linnaeus and his contemporaries founded biogeography—a science that aims to document and understand spatial patterns of biological diversity—by studying organisms visible to the naked eye. Recent advances in our ability to quantify and visualize microbial diversity in natural environments have prompted a new era of microbial exploration, one that builds upon plant and animal biogeography surveys initiated roughly 250 years ago (*1*). These new explorations have already radically changed thinking in ecology and evolution and upset the hierarchical taxonomic structure that Linnaeus proposed (*2*). As microbiologists increasingly focus on biogeographical questions, textbook placeholders such as "microbial biogeography is poorly known and rarely discussed" (*3*) will become obsolete.

A long-held concept in microbial ecology is that microorganisms are dispersed globally and able to proliferate in any habitat with suitable environmental conditions. This concept was prompted by Martinus Willem Beijerinck and concisely summarized by Lourens Gerhard Marinus Baas Becking in the widely referenced quote, "everything is everywhere, but the environment selects" (*4*). Seminal notions of ubiquitous dispersal and environmental determinism are not unique to microbiology. Linnaeus, for example, wrote that "the great Artificer of Nature has provided that every seed shall find its proper soil, and be equally dispersed over the surface of the globe" (*1*). The development of molecular approaches has allowed a more comprehensive view of microbial diversity than can be developed even with the aided eye, showing that like plant and animal distributions, microbial distributions can be the result of both deterministic (environmental) and stochastic (dispersal) processes [reviewed in (*5*)].

As with macroorganism biogeography, microbial biogeography initially adopted a taxonomic approach, focusing on sequence signatures to identify groups of microorganisms. These studies revealed classic patterns such as the species-area relationship and isolation by distance (*5*). Interpreting taxonomic patterns in terms of how they affect the function of a population or community is especially difficult in microorganisms, where a broad range of functional variation may occur among similar organisms (e.g., organisms with the same 16S rRNA sequence).

For macro- and microorganisms alike, there is growing interest in the biogeography of functional traits, characteristics of an organism that are linked with its fitness or performance (*6*). The study of biogeography includes the study of patterns in space, in time, and along environmental gradients. Such patterns in the distribution of traits can be used to understand complex phenomena, including why organisms live where they do, how many taxa can coexist in a place, and how they will respond to environmental change. Although plants have been the focal group in this emerging research area, recent advances in environmental molecular biology such as genomics, proteomics, transcriptomics, and metabolomics place microbial ecology in a unique position to move trait-based biogeography forward.

Trait-Based Biogeography: A Macroorganism Perspective

Trait-based approaches to biogeography have been used since the pioneering work of Andreas F. W. Schimper more than a century ago (*7*). Although tending to wax and wane in favor over time, there has been a resurgence of interest in trait-based methodologies since the mid-1980s (*6*, *8*). Here, we discuss some examples of plant trait-based research, focusing on applications likely relevant to both plant and microbial ecology.

An emergent theme in trait-based research is the identification of ecological strategies, suites of covarying ecological traits. The study of ecological strategies has been fundamental to the development of plant and animal ecology, and there is growing interest among microbial ecologists as well (*9*). Examining the slope, intercept, and correlation strength of relationships among traits provides insight into the nature of ecological strategies, including the underlying costs and benefits of different trait combinations. Quantification of trait variation with site properties such as climate, for example, is central to understanding how vegetation properties shift along geographical gradients, and thus for predicting habitat boundaries under changing land-use and warming scenarios. A similar approach has been suggested for microorganisms (*9*) and could be useful for predicting how microbial properties respond to environmental change.

Figure 1A illustrates how the relationship between two ecologically important plant traits—leaf life span and leaf mass per area—shifts with climate (*10*). This trait relationship is part of a fundamental ecological strategy known as the "leaf economics spectrum" that ranges from organisms with cheaply constructed, thin, and short-lived leaves to those with costly, thick, and long-lived leaves. The relationship is modulated by site climate, such that organisms at drier sites typically achieve shorter leaf life spans at a given leaf mass per area.

Community "assembly rules" was a concept formulated to understand why organisms live where they do and in what combinations. It has been used to understand how plant and animal communities change through time and to predict

[1]Center for Ecology and Evolutionary Biology, Department of Biology, University of Oregon, Eugene, Oregon 97403, USA. [2]Department of Microbiology, University of Illinois Urbana-Champaign, Urbana, Illinois 61801, USA

*To whom correspondence should be addressed. E-mail: jlgreen@uoregon.edu

the response of communities to environmental change. This concept has also grown in popularity among microbial ecologists (*11*). Weiher and Keddy (*12*) redefined the assembly-rule paradigm by introducing the idea that traits, not taxon names, are the fundamental units of biodiversity and biogeography. They argued that a trait-based approach would alleviate reliance on "murky taxonomy" and introduced a conceptual model for trait-based community assembly. Organisms with similar traits will share similar niche requirements, which will result in them "sorting" into similar environments. At the same time, environmental sorting is expected to lead to competitive exclusion among organisms with very similar traits. An informative counter-perspective is the neutral theory (*13*), a community-assembly model that often makes realistic biogeography predictions despite explicitly ignoring trait variation.

Empirical patterns of trait variation across environmental, spatial, and temporal gradients are central to testing hypotheses arising from community assembly models. In a study of spatial variation in traits (Fig. 1B), Swensen and Enquist (*14*) showed that the mean and variance in community aggregated wood density was negatively correlated with latitude. Their results comply with Weiher and Keddy's model and suggest that on a global scale, there is strong environmental filtering of wood density values along the latitudinal gradient, whereas on local scales competition becomes increasingly important toward the equator. Shipley and colleagues (*15*) integrated Weiher and Keddy's qualitative model with tools borrowed from statistical mechanics to develop a quantitative method for predicting biodiversity patterns across environments. They tested their model using trait variation data collected from abandoned vineyards over successional time (Fig. 1C) and predicted the relative abundance of species at each successional age with great accuracy. Similar questions have been asked using phylogenetic analyses of microbial taxa (*16*); a focus on traits requires fewer assumptions and more directly addresses microbial properties important to ecosystem function. Such an approach could be valuable for predicting the temporal dynamics of microbial communities and the consequences of environmental change.

Trait-Based Biogeography: A Microbial Perspective

The advancement of environmental molecular biology, in combination with laboratory studies of microbial physiology and metabolism, has provided novel tools for identifying trait patterns in microorganisms on a scale not yet possible for studies of plants and animals. However, we currently know little (relative to our understanding for plants and animals) regarding the distribution of microbial functional traits in nature. Our lack of understanding is due in part to the extraordinary metabolic and physiologic diversity of microbes, our inability to culture many microorganisms, and the common practice of reducing trait variation (for example, by studying single isolates in culture). The value of studying variation in traits is rapidly becoming apparent to many microbiologists as new patterns of variation are discovered, even within groups once thought to be relatively homogeneous. New tools have become available that allow the study of microbial traits in their natural contexts without necessarily requiring the culturing of the individual organisms.

One of the biggest challenges facing microbiologists interested in functional trait is how to identify them. As for plants and animals, a microbial functional trait is a measurable microbial characteristic that is linked to fitness or performance. Among the many possible types of traits that could fit this definition are those that directly contribute to an organism's life history (its lifetime pattern of growth, reproduction, and differentiation), traits that mediate the interactions among organisms (e.g., microbial resistance to viruses), and traits that contribute to the storage, transfer, and organization of information (e.g., genome size, content, and mutation rate). The challenge is to determine which characteristics of microorganisms are linked most closely to fitness or performance. To date, this has primarily been accomplished by studying trait variation among cultured isolates. Microbiologists currently have a detailed understanding of only a small proportion of microbial diversity, primarily those taxa with a long history of laboratory cultivation and study. Microbiologists have comprehensive information on trait variation in the field for only a handful of these, most notably the phototrophs *Prochlorococcus* and *Synecococcus* (*17*).

For organisms that are not as easy to culture, it is possible from a general understanding of microbial biochemical pathways, energetics, and the mechanisms of metabolic regulation to choose traits that are likely to be linked to fitness. For example, based on decades of research on archaeal isolates, microbiologists have a highly developed understanding of the structure and function of bacteriorhodopsin (a light-sensitive membrane protein capable of generating a proton gradient that can be harvested for energy). This knowledge was used to identify a type of rhodopsin in bacteria—proteorhodopsin—before the existence of any cultured bacterial isolates (*18*).

Functional traits can also be identified by engineering genetic variation identified from environmental genomic surveys into non-native cultured hosts. Genetic variants of the gene that codes for proteorhodopsin have been inserted into *Escherichia coli*, resulting in a photochemical reac-

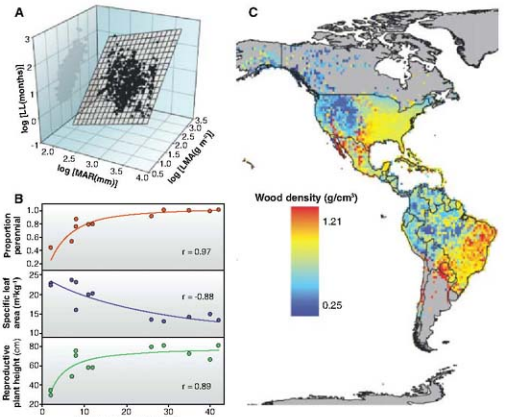


Fig. 1. Examples of trait-based biogeography for macroorganisms. **(A)** Covariation in leaf traits [leaf life span (LL) and leaf mass per area (LMA)] as a function of climate [mean annual rainfall (MAR)]. Data plotted for species sampled worldwide [modified from (*10*)]. **(B)** Variation in community-aggregated plant traits over successional time. Data sampled from vineyards in France after abandonment [modified from (*15*)]. **(C)** Variation in community-aggregated wood density with latitude. Data plotted for mean wood density of all species in 1° grid cells [modified from (*14*)].

tion cycle characteristic of archaeal rhodopsins (*18*). The presence of proteorhodopsin is now considered a putative functional trait in bacteria, and variants of proteorhodopsin with different spectral properties have been shown to exhibit spatial and temporal patterns in marine environments (*19*).

Another way to identify putative functional traits from environmental samples is to study relationships among traits. For example, Ram *et al.* (*20*) used environmental proteomics to identify the prevalence of proteins involved in cobalamin and heme synthesis that are abundantly expressed in combination with cytochrome$_{579}$, a molecule known to be important in iron oxidation in acid mine drainage biofilms. The fact that these two proteins covary with a protein likely linked to fitness suggests that their presence may be a functional trait.

Table 1 illustrates some examples of putative microbial functional traits that have been studied in a biogeographical context. These examples underscore the value of culture-dependent approaches combined with high-throughput technologies that allow for the sampling of genes, proteins, expression patterns, and metabolites in situ. Below we highlight several studies that have revealed patterns of microbial functional trait variation across time, space, and environmental gradients.

Ribosomal Genes and Ecological Succession

One putative functional trait that has been identified for several macro- and microorganisms is ribosomal gene copy number. Ribosomal copy number is one of a suite of life history traits associated with general ecological strategies in bacteria (*9*), and it may play an important role in microbial community assembly (*21*). The copy number of ribosomal DNA (rDNA) genes is related to ribosome content and has been linked in microorganisms to maximum growth rate and response time after an increase in resources (*21*). There also appears to be a competitive cost associated with possessing multiple gene copies under low resource conditions (*21*), which suggests a life history trade-off.

Shrestha *et al.* (*22*) tracked microbial ribosomal copy number in a rice paddy soil during succession after flooding. They observed that isolates from early successional soils (after 1 day of flooding) had significantly higher ribosomal gene copy numbers than those from later successional soils (after 70 days of flooding). Isolates from early successional soils also formed colonies on solid media more rapidly than those from late successional soils (Fig. 2A), which suggests a faster response time to an increase in nutrients. Succession is one of several temporal patterns that have been studied extensively in macroorganisms by biogeographers, and patterns in functional traits and ecological strategies have been documented (Fig. 1B).

Ribosomal gene copy number has the potential to have wide-ranging effects on ecosystem function through its effects on cellular stoichiometry and food web interactions (*23*), and understanding the biogeography of ribosomal gene copy number could be important for predicting the response of ecosystems to environmental change. It is difficult to measure microbial rDNA copy number in situ, and to date most measures have been made on laboratory isolates from environmental samples (*22*) or microcolonies grown on filters placed in environmental samples (*24*). Advances in quantitative polymerase chain reaction (qPCR), environmental genomics, and single-cell genomics may soon allow the quantification of ribosomal copy number in situ, on both a community and individual level.

Metabologeography

Metabolomics can be used to characterize microorganisms across space by analysis of the small molecules that are the intermediates and products of metabolism and which leave chemical fingerprints on specific cellular processes. Variation in metabolic profiles not only can be used as an indirect measure of variation in microbial functional traits but also can indicate interactions among individuals, including cell-cell communication and competition (for example, through antibiotic production).

Rossello-Mora *et al.* (*25*) sampled isolates of the bacterium *Salinibacter ruber*, from three regions: the Mediterranean, the Peruvian Andes, and the Canary Islands. Genetic methods did not reveal biogeographic patterns among isolates. When high-resolution mass spectrometry was used to identify and quantify metabolites, biogeographic patterns among the isolates were evident (Fig. 2B). Molecules associated with the cell envelope, such as sulfonolipids, left the strongest signature. The differences seen between isolates from different geographic regions were primarily quantitative rather than qualitative, implying that transcriptional and/or posttranscriptional regulation was involved. To make a similar analysis at the community level without culture requires metabolite assays on environmental samples, and to our knowledge this has not yet been done. However, protein analysis at the community level has been successful for environmental samples (*20*), and the field of environmental metabolomics is rapidly growing.

Environmental Gradients and Genomics

One approach to studying microbial traits without relying on culture is to make spatially explicit use of sequences from environmental genomics studies. Gene sequences can be used to assign putative physiological traits through sequence similarity to genes of known function by using annotation databases, such as the Kyoto Encyclopedia of Genes and Genomes, Clusters of Orthologous Groups, and SEED subsystems. DeLong *et al.* (*26*) used this approach to analyze samples of microbial plankton from seven depths in the Pacific Ocean. They prepared community DNA fosmid libraries, end-sequenced ~5000 fosmids from each depth, and compared the distribution of genes with depth. From this distribution of genes in space, they were able to identify a number of patterns in putative functional traits (Fig. 2C). As expected, sequences associated with photosynthesis were more common in samples from the photic zone, as were sequences associated with motility. This is consistent with the hypothesis that heterotrophic bacteria in the photic zone may actively compete for resources by swimming from particle to particle. In contrast, sequences associated with surface growth, for example, pilus synthesis, protein export, and polysaccharide and antibiotic synthesis genes, were more prevalent in deeper samples, suggesting an increased importance of a relatively sessile lifestyle at these depths. This approach was able to document, at the community level, patterns in the distribution of single putative functional traits and patterns in suites of covarying traits (e.g., those associated with photosynthesis, motility,

Table 1. Selected examples of microbial functional traits.

Trait	Potential consequence for fitness or performance	Measure	Example
Gene sequence identity	Protein function	PCR amplification	(*19*)
Ribosomal copy number	Growth rate, response time	Southern hybridization	(*22*)
Metabolite content	Survival	Mass spectrometry	(*25*)
Gene content	Metabolic potential, life history variation	Genomics	(*26*)
Genome size	Metabolic potential, life history variation	Genomics	(*29*)
Environmental resistance (antibiotic, viral, radiation, etc.)	Survival	PCR amplification, laboratory screening of isolates	(*30*)
Transcript abundance	Metabolic switching	Reverse transcription PCR, qPCR	(*31*)
Mutation rate	Adaptability	Laboratory assays of isolates	(*32*)
Cell size	Metabolic rate	Flow cytometry	(*33*)

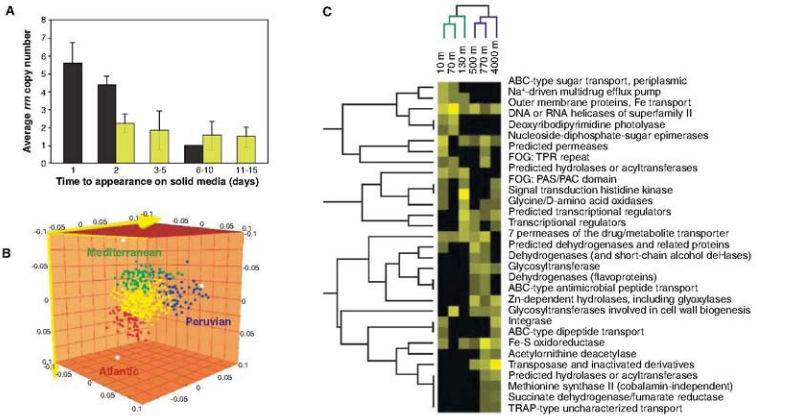


Fig. 2. Examples of microbial trait-based biogeography. **(A)** Mean numbers of rDNA operon copies in early successional (black bars) and late successional (yellow bars) bacterial taxa isolated from rice paddy soil [modified from (*22*)]. **(B)** Variation in the metabolite composition of *Salinibacter ruber* isolates from different geographic regions. The metabolites having a high correlation with geographic origin are noted with a corresponding color; the nondiscriminating metabolites are represented in yellow [modified from (*25*)]. **(C)** Variation in the relative abundance of protein categories (inferred via the COG database from environmental DNA sequences) along a depth transect in the Pacific Ocean [modified from (*26*)].

and surface growth) that may indicate the spatial patterning of ecological strategies.

Outlook

Trait-based approaches to microbial biodiversity and biogeography offer the promise of advancing ecological theory and predicting responses to environmental change. Recent theoretical advances in plant sciences using a trait-based framework have provided a mechanistic basis for quantifying the link between functional trait variation and ecosystem processes (*15*, *27*, *28*). A focus on traits that are common to macroorganisms and microorganisms provides a common currency with which to develop and test theory across life's domains. The next step is to revisit these theories in the context of microbial systems.

The historical emphasis on taxonomy-based conservation has been mired in the argument of functional redundancy, which assumes that taxa are functionally interchangeable. This idea has been especially influential in microbial ecology, resulting in the assumptions that neither a loss in microbial taxonomic diversity nor a turnover in microbial community composition will have consequences for microbial-mediated processes, because many different microbial taxa can mediate the same process. In addition, because microorganisms are assumed to evolve rapidly, microbial taxa distributions have been assumed to be of little value in predicting the response of microbial communities to environmental change. A trait-based approach will recast this debate to better understand the importance of specific suites of microbial functional traits in the environment.

Schimper's seminal text on plant geography (*7*) is prefaced with the opening statement, "the time is not far distant when all species of plants and their geographical distribution will be known. The objects of geographical botany will not, however, be attained, as is often assumed, but a foundation merely will have been laid on which science can construct a larger edifice. The essential aim of geographical botany will then be an inquiry into the causes of differences existing among the various floras." Building on Linnaeus' taxonomic biogeography, Schimper adopted an ecological approach, summarizing in elaborate detail the form, structure, physiology, and life history of plants around the world. Schimper's vision gives us a universally applicable framework for trait biogeography to which we can now add microorganisms.

References and Notes

1. C. Linnaeus, *Amoenitates Academicae* **2**, 17 (1781).
2. C. R. Woese, O. Kandler, M. L. Wheelis, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 4576 (1990).
3. M. V. Lomolino, B. R. Riddle, J. H. Brown, *Biogeography* (Sinauer Associates, Inc., Sunderland, Massachusetts, ed. 3, 2006).
4. L. G. M. Bass Becking, *Geobiologie of Inleiding tot de Milieukunde* (W. P. van Stockum and Zoon, The Hague, 1934).
5. J. B. H. Martiny *et al.*, *Nat. Rev. Microbiol.* **4**, 102 (2006).
6. B. J. McGill, B. J. Enquist, E. Weiher, M. Westoby, *Trends Ecol. Evol.* **21**, 178 (2006).
7. A. F. W. Schimper, *Plant Geography Upon a Physiological Basis* (Clarendon Press, Oxford, 1903) [translation of A. F. W. Schimper, *Pflanzengeographie Auf Physiologischer Grundlage* (G. Fischer, Jena, Germany, 1898)].
8. M. Westoby, I. J. Wright, *Trends Ecol. Evol.* **21**, 261 (2006).
9. N. Fierer, M. A. Bradford, R. B. Jackson, *Ecology* **88**, 1354 (2007).
10. I. J. Wright *et al.*, *Nature* **428**, 821 (2004).
11. L. DeMeIfsen, P. B. Eckburg, E. M. Bik, D. A. Relman, *Trends Ecol. Evol.* **21**, 517 (2006).
12. E. Weiher, P. A. Keddy, *Oikos* **74**, 159 (1995).
13. S. P. Hubbell, *Funct. Ecol.* **19**, 166 (2005).
14. N. G. Swenson, B. J. Enquist, *Am. J. Bot.* **94**, 451 (2007).
15. B. Shipley, D. Vile, É. Garnier, *Science* **314**, 812 (2006).
16. M. C. Horner-Devine, B. J. M. Bohannan, *Ecology* **87**, 5100 (2006).
17. Z. I. Johnson *et al.*, *Science* **311**, 1737 (2006).
18. O. Beja *et al.*, *Science* **289**, 1902 (2000).
19. G. Sabehi *et al.*, *ISME J.* **1**, 48 (2007).
20. R. J. Ram *et al.*, *Science* **308**, 1915 (2005).
21. B. S. Stevenson, T. M. Schmidt, *Appl. Environ. Microbiol.* **70**, 6670 (2004).
22. P. M. Shrestha, M. Noll, W. Liesack, *Environ. Microbiol.* **9**, 2464 (2007).
23. L. J. Weider *et al.*, *Annu. Rev. Ecol. Evol. Syst.* **36**, 219 (2005).
24. S. J. Binnerup *et al.*, *FEMS Microbiol. Ecol.* **37**, 231 (2001).
25. R. Rossello-Mora *et al.*, *ISME J.* **2**, 242 (2008).
26. E. F. DeLong *et al.*, *Science* **311**, 496 (2006).
27. V. M. Savage, C. T. Webb, J. Norberg, *J. Theor. Biol.* **247**, 213 (2007).

28. B. J. Enquist *et al.*, *Nature* **449**, 218 (2007).
29. K. T. Konstantinidis, J. M. Tiedje, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 3160 (2004).
30. R. S. Singer, M. P. Ward, G. Maldonado, *Not. Rev. Microbiol.* **4**, 943 (2006).
31. A.-S. Steunou *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 2398 (2006).
32. A. Giraud *et al.*, *Science* **291**, 2606 (2001).
33. A. Rinaldo, A. Maritan, K. K. Cavender-Bares, S. W. Chisholm, *Proc. R. Soc. London B* **269**, 2051 (2002).
34. We thank I. Wright, M. Westoby, K. Milferstedt, B. McGill, N. Martinez, E. Berlow, and the University of Oregon Microbial Ecology Journal Club for valuable feedback; J. Harte for encouraging us to think about microbial biodiversity in new ways; and N. Swensen and E. Garnier for providing figures. J.L.G., B.J.M.B., and R.J.W. acknowledge the support of the National Science Foundation. J.L.G. also acknowledges support from the Gordon and Betty Moore Foundation.

10.1126/science.1153475

REVIEW

Microbial Ecology of Ocean Biogeochemistry: A Community Perspective

Suzanne L. Strom

The oceans harbor a tremendous diversity of marine microbes. Different functional groups of bacteria, archaea, and protists arise from this diversity to dominate various habitats and drive globally important biogeochemical cycles. Explanations for the distribution of microbial taxa and their associated activity often focus on resource availability and abiotic conditions. However, the continual reshaping of communities by mortality, allelopathy, symbiosis, and other processes shows that community interactions exert strong selective pressure on marine microbes. Deeper exploration of microbial interactions is now possible via molecular prospecting and taxon-specific experimental approaches. A holistic outlook that encompasses the full array of selective pressures on individuals will help elucidate the maintenance of microbial diversity and the regulation of biogeochemical reactions by planktonic communities.

When we try to pick out anything by itself, we find it hitched to everything else in the universe.
John Muir (1911)

Marine microbes are fundamental regulators of biogeochemical cycles [see the accompanying review by Falkowski *et al.* (*1*)]. While acquiring resources for metabolism and growth, archaea, bacteria, and protists transform C-, N-, P- and S-containing compounds in ways that affect their availability for biological production and their influence on Earth's climate. Questions about the relationships between plankton ecology and these transformations are at the heart of much ocean research and have existed since the investigations of Brandt, Lohmann, and others a century ago [as described in (*2*)]. Yet, despite a burgeoning toolbox of methods, most research into ocean biogeochemistry and associated microbial ecology is still framed in terms of "bottom-up" considerations (i.e., how do resources and abiotic conditions drive the distribution and function of microbes?). Ecological considerations and recent genomic findings demonstrate that a broader view is necessary. Agents of mortality, including grazers and lytic viruses, must exert strong selective pressure on populations of marine microbes. Additional interactions such as allelopathy and symbiosis also influence community structure and function. Further, these processes interact with resource availability in numerous ways so that separate bottom-up versus "top-down" approaches to studying these ecosystems are of limited conceptual or experimental use. Marine microbes are truly hitched to everything else in the ocean universe. Thus, a community perspective is essential for understanding the distribution and function of microorganisms in Earth's oceans. By applying the comucopia of new ocean research methods to questions of whole-community structure and function, we will gain insights into the regulation of biogeochemical cycling. A community perspective will also aid our understanding of the sources and functions of the vast genomic diversity housed in the oceans' microbes.

The Broader View: Ecological Considerations

Owing to physical and chemical constraints, nearly all primary and most secondary production in the oceans is performed by microbes. Little energy storage or accumulation of structural material is possible in a unicell; hence, in striking contrast to terrestrial ecosystems, almost no marine production accumulates as biomass. Except during bloom events, nearly all marine microbial production "turns over" in days to weeks through various mortality processes (*3*). In ocean regions remote from land, microbial mortality is mainly due to grazing by protists (*4*, *5*); in other areas, disease, viral lysis, grazing by larger zooplankton, and perhaps microbial senescence can be substantial (*6*, *7*). Cumulative mortality in approximate equivalence with growth leads to a continual reshaping of the community at a fundamental level and a situation in which avoidance of mortality is tantamount to resource-based growth as a selective pressure on individual microbes (Fig. 1). From these considerations, one would predict that adaptations reducing mortality (e.g., escape behavior, defenses, resistance to infection, camouflage) must constitute a substantial part of the genotypic and phenotypic repertoire of marine microbes.

The Broader View: Genomic Considerations

Sequencing of cultured marine microbes and metagenomic surveys of natural environments are revealing genes related to community interactions. Using shotgun sequencing of fosmid clone termini, DeLong and co-workers (*8*) assessed the depth distribution of microbial genes in the North Pacific subtropical gyre. Deep-water sequences were enriched in genes for antibiotic synthesis, which are hypothesized to play a role in structuring particle-associated communities. The cyanobacterium *Synechococcus* is a major prokaryotic contributor to ocean primary production, and genomes of several isolates have been sequenced. Both Sargasso Sea and California Current isolates contain several different polysaccharide biosynthesis genes; the authors speculate that these allow changes in cell-surface characteristics, providing camouflage from or resistance to phage or predator attack (*9*). *Prochlorococcus* is another important genus of photosynthetic prokaryotes, one in which the links between genetics and ecophysiology are particularly well characterized. A survey of *Prochlorococcus* "ecotypes" in the Atlantic Ocean found that environmental gradients in light, temperature, and nutrient availability explained considerably less than half of the variability in spatial distribution for all but one ecotype (*10*). Unexplained aspects of distribution hinted at important but poorly characterized differences in competitive ability and susceptibility to grazing and viral lysis.

In the following sections, I present examples of how community processes can drive biogeochemical cycling of Si, C, and N, as well as the climatically important gas dimethyl sulfide (DMS). I then outline two research areas in which further investigation promises to elucidate mechanistic links between community ecology and biogeochemistry.

Paradigms Lost

Current paradigms, largely revolving around resource acquisition and abiotic conditions, are often insufficient to explain major patterns of functional group dominance in the sea [see the accompanying

Shannon Point Marine Center, Western Washington University, Anacortes, WA 98221, USA. E-mail: Suzanne.Strom@wwu.edu

review by Green *et al.* (*11*)]. For example, in the absence of chronic iron limitation, the spring primary production increase in temperate and high-latitude seas is largely manifested as blooms of diatoms. There is little evidence from experimental laboratory-based work that diatoms have consistently higher growth rates than photosynthetic flagellates under the cold, high-nutrient conditions of early spring (*12*), nor are they consistently more tolerant of turbulence (*13*). Superior nutrient uptake and storage capacity have also been invoked as determinants of diatom success in the spring bloom niche (*14*). However, during the inception and increase of the spring bloom, nutrient concentrations far surpass those limiting to uptake or growth. Thus, diatom blooms cannot be explained solely as the outcome of superior environmental tolerance or "simple" resource competition among photosynthetic protists (although blooms are often modeled this way). We do know that the diatom species that dominate blooms experience less grazing mortality than do co-occurring species (*15*, *16*). The reasons are poorly understood, and the existing hypotheses are not yet strongly substantiated. One possibility is that blooming diatoms have grazing-resistant morphologies; another is that they produce defensive chemicals (*17*, *18*). There may also be a mismatch between temperature optima for growth of diatoms relative to growth of their potential predators (*12*). Allelopathic inhibition of competitors by diatoms has been hypothesized as a contributing factor (*19*). The need for a mechanistic and hence a predictive understanding of diatom blooms is a priority because of their substantial biogeochemical and ecological consequences, including their role in ocean C and Si cycling, as well as their importance in food webs supporting large benthic and pelagic animals.

For other major biogeochemical processes, including global rates of N_2 fixation and denitrification, we have very little knowledge of how communities might regulate the abundance and biogeochemical transformation rates of the relevant microbes. Denitrifiers inhabit low-oxygen zones in the ocean's interior. Observations from the Black Sea and European fjords show that protistan grazers of bacteria, such as ciliates and heterotrophic flagellates, also inhabit low-oxygen waters. We also know that microaerobic and anaerobic protists can form consortia with bacteria (*20*, *21*). However, the role of these predatory and symbiotic relationships in regulating the abundance and activity of denitrifiers is largely unknown.

Volatile Communities

Biogeochemical processes involving ocean-atmosphere S exchange are strongly dependent on community composition and function. The volatile compound DMS is produced by both prokaryotic and eukaryotic marine microbes through enzymatic cleavage of dimethyl sulfoniopropionate (DMSP). Fluxes of DMS from the ocean to the atmosphere directly influence Earth's climate, because DMS molecules can act as cloud condensation nuclei, increasing cloud cover over the ocean. DMS fluxes have been remarkably difficult to predict from environmental variables alone, because community interactions play a role in nearly all aspects of DMS production. In addition to an abundance of DMSP-containing microbes, large DMS fluxes can only be generated if there are high rates of processes promoting DMSP cleavage combined with low rates of bacterial DMSP demethylation; the latter converts the molecule to products other than DMS.

Communities determine DMS fluxes by several processes (Fig. 2), including preferential grazing on prey that contain low levels of DMSP or lyase. Extensive interstrain variability in these properties has been found (*22*, *23*), which may be related to differential allocation of cellular resources to resource acquisition, defense, signaling, and other survival needs. Although poorly understood, preferential grazing could also select for either demethylating or DMS-producing bacteria. Released DMSP is a chemical signal that can reduce grazing, potentially promoting blooms, and viral lysis can cause release of DMS from host cells (*24*, *25*). Although still incompletely understood, this is one of the most thoroughly investigated microbial systems of those involved in a major biogeochemical transformation. Data collected to date demonstrate that a knowledge of community processes will be essential for a mechanistic and predictive understanding of DMS fluxes. There is little doubt that community ecology will prove equally important for understanding other biogeochemical cycles.

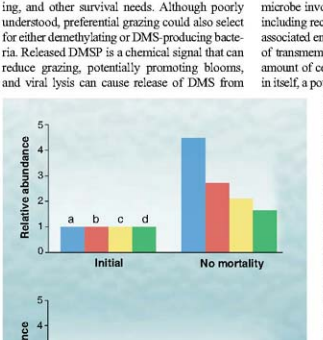


Fig. 1. Hypothetical microbial community with four major taxonomic groups (indicated by the lowercase letters and associated colors), present initially (**top left**) at equal abundances. With no mortality (**top right**), the fastest-growing taxon ("a") dominates the community after 5 days of growth. (**Bottom left**) Equal rates of mortality on all taxa yield reduced abundances—hence, potentially reduced biogeochemical activity—due to generalized removal of microbes while they are growing. (**Bottom right**) Unequal rates of mortality lead to shifts in species dominance and fundamental restructuring of the community. In this case, the two slowest-growing taxa ("c" and "d") dominate after 5 days. Growth rates [k (days^{-1})]: a = 0.30, b = 0.20, c = 0.15, d = 0.10; mortality rates [g (days^{-1})]: a = 0.35, b = 0.13, c = 0.02, d = 0.01; abundance (A) at 5 days = $A_{initial}e^{(k-g)5}$. Data taken from (*38*).

Cell Surfaces: A Focus of Selective Pressures

The role of the cell surface in simultaneously mediating resource acquisition and defending or camouflaging the cell against agents of mortality is poorly understood for free-living microorganisms. Features of pathogen cell surfaces are better known and might offer some insights into universal mechanisms. Every resource acquired by a microbe involves some aspect of the cell surface, including receptors, transporters, and cell surface–associated enzymes. Given the high ion specificity of transmembrane pumps and channels (*26*), the amount of cell surface area for such structures is, in itself, a potentially limiting resource. How, then, does the constellation of resource uptake requirements affect the landscape of the individual cell surface? It is this landscape that many grazers perceive and respond to during feeding. We know that alteration of prey cell surfaces by lectin binding or coating with various organics affects the feeding rates of flagellates (*27*, *28*). Marine *Synechococcus* genomes contain multiple polysaccharide synthesis genes (*9*, *29*); analogous to the antigenic variation of pathogens to avoid host-mediated phagocytosis (*30*), these genes may allow cell-surface changes in response to grazing pressure. Host specificity of viral infection may also depend on specific cell-surface oligosaccharides (*31*). These and other observations indicate that mortality operates in part through recognition (or lack thereof) of cell-surface properties, which in turn may be influenced by resource acquisition. A research focus on microbial cell surfaces is likely to elucidate how these organisms successfully adapt to the suite of selective pressures they face.

Allelopathy

Allelopathic interactions, in which microbes actively inhibit their competitors, are poorly understood for marine microbial communities. Allelopathy has long been invoked as a mechanism promoting blooms, although distinguishing inhibition from resource competition can be experi-

mentally challenging (*19*). Production of lytic agents has been demonstrated for both prokaryotic and eukaryotic marine microbes and interpreted as a mechanism for simultaneously reducing competition or predation and obtaining dissolved organic nutrients [i.e., "kill and eat your predator" (*32*)]. Allelopathy may be more widespread than anticipated and may encompass tactics including the release of reactive oxygen species and antibiotics (*33*, *34*). Sublethal allelopathy could be particularly effective in shaping community structure and function by altering growth and loss rates. For example, microbes are known to release nontoxic signaling compounds that alter the behavior of conspecifics, competitors, and/or predators (*35*, *36*). The release of siderophores that bind up iron and other trace nutrients in a form unavailable to competitors could also be seen as a form of allelopathic inhibition (*37*).

Research Prospects

Marine microbes drive globally important biogeochemical cycles, and we are in the midst of an explosion of data concerning the genomic diversity that underlies this microbial activity. At the same time, there is growing evidence that community interactions—including various sources of mortality, allelopathy, symbiosis, and others—exert strong selective pressure on marine microbes. Marine microbial distribution and diversity cannot be understood without considering the entire suite of selective pressures on individual microbes. By the same token, study of community processes is essential for an understanding of ecosystem function, including important biogeochemical transformations. Progress will require that genomic and other surveys be conducted and interpreted with an eye for molecules that underlie not only environmental tolerance and resource acquisition but also cell-cell and population interactions of various types. New molecular techniques are beginning to be coupled with community-level experimentation and hypothesis testing; there is ample scope for more research of this type, especially that which targets key functional groups of marine microbes. Implementing these suggestions will require collaborations between scientists with historically divergent views of marine microbes, including community ecologists, geneticists, and oceanographers. The resulting holistic approach to how we conceptualize, observe, and experiment with microbial communities will be required to achieve a synthesis of microbial ecology and biogeochemistry.

CREDIT: K. SUTTLE/SCIENCE

References and Notes

1. P. G. Falkowski, T. Fenchel, E. F. DeLong, *Science* **320**, 1034 (2008).
2. E. L. Mills, *Biological Oceanography: An Early History, 1870–1960* (Cornell Univ. Press, Ithaca, NY, 1989).
3. A. R. Longhurst, *Ecological Geography of the Sea* (Academic, San Diego, CA, 1998).
4. S. L. Strom, in *Microbial Ecology of the Oceans*, D. L. Kirchman, Ed. (Wiley-Liss, New York, 2000), pp. 351–386.
5. A. Calbet, M. R. Landry, *Limnol. Oceanogr.* **49**, 51 (2004).
6. C. A. Suttle, *Nat. Rev. Microbiol.* **5**, 801 (2007).
7. K. D. Bidle, P. G. Falkowski, *Nat. Rev. Microbiol.* **2**, 643 (2004).
8. E. F. DeLong *et al.*, *Science* **311**, 496 (2006).
9. B. Palenik *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 13555 (2006).
10. Z. I. Johnson *et al.*, *Science* **311**, 1737 (2006).
11. J. L. Green, B. J. M. Bohannan, R. J. Whitaker, *Science* **320**, 1039 (2008).
12. J. M. Rose, D. A. Caron, *Limnol. Oceanogr.* **52**, 886 (2007).
13. T. J. Smayda, *Harmful Algae* **1**, 95 (2002).
14. E. Litchman, C. A. Klausmeier, O. M. Schofield, P. G. Falkowski, *Ecol. Lett.* **10**, 1170 (2007).
15. S. L. Strom, E. L. Macri, M. B. Olson, *Limnol. Oceanogr.* **52**, 1480 (2007).
16. P. Assmy, J. Henjes, C. Klaas, V. Smetacek, *Deep-Sea Res. I* **54**, 340 (2007).
17. C. E. Hamm *et al.*, *Nature* **421**, 841 (2003).
18. G. Pohnert, *ChemBioChem* **6**, 946 (2005).
19. C. Legrand, K. Rengefors, G. O. Fistarol, E. Granéli, *Phycologia* **42**, 406 (2003).
20. T. Fenchel *et al.*, *Ophelia* **43**, 45 (1995).
21. M. V. Zubkov, A. F. Sazhin, M. V. Flint, *FEMS Microbiol. Lett.* **101**, 245 (1992).
22. S. L. Strom *et al.*, *Limnol. Oceanogr.* **48**, 217 (2003).
23. M. Steinke, C. Daniel, G. O. Kirst, in *Biological and Environmental Chemistry of DMSP and Related Sulfonium Compounds*, R. P. Kiene, P. Visscher, M. Keller, G. Kirst, Eds. (Plenum, New York, 1996), pp. 317–324.
24. W. H. Wilson *et al.*, *J. Mar. Biol. Assoc. U.K.* **82**, 369 (2002).
25. S. L. Strom, G. V. Wolfe, A. Slajer, S. Lambert, J. Clough, *Limnol. Oceanogr.* **48**, 230 (2003).
26. E. Gouaux, R. MacKinnon, *Science* **310**, 1461 (2005).
27. E. C. Wootton *et al.*, *Environ. Microbiol.* **9**, 216 (2007).
28. C. Matz, J. Boenigk, H. Arndt, K. Jurgens, *Aquat. Microb. Ecol.* **27**, 137 (2002).
29. B. Palenik *et al.*, *Nature* **424**, 1037 (2003).
30. P. Henneke, D. T. Golenbock, *J. Exp. Med.* **199**, 1 (2004).
31. H. Mitamoto, Y. Tomaru, Y. Takao, Y. Shirai, K. Nagasaki, *J. Virol.* **81**, 1372 (2007).
32. U. Tillmann, *Aquat. Microb. Ecol.* **32**, 73 (2003).
33. J. Marshall, M. Salas, O. Tatsuya, G. M. Hallegraeff, *Mar. Biol. (Berlin)* **147**, 533 (2005).
34. R. A. Long, F. Azam, *Appl. Environ. Microbiol.* **67**, 4975 (2001).
35. S. L. Strom, G. V. Wolfe, K. B. Bright, *Aquat. Microb. Ecol.* **47**, 107 (2007).
36. A. Vardi *et al.*, *PLoS Biol.* **4**, e60 (2006).
37. D. A. Hutchins, A. E. Witter, A. Butler, G. W. Luther III, *Nature* **400**, 858 (1999).
38. B. W. Frost, in *The Physiological Ecology of Phytoplankton*, I. Morris, Ed. (Univ. of California Press, Berkeley, CA, 1980), vol. 7, pp. 465–491.
39. G. V. Wolfe, M. Steinke, *Limnol. Oceanogr.* **41**, 1151 (1996).
40. P. A. Matrai, M. D. Keller, *Cont. Shelf Res.* **13**, 831 (1993).
41. J. Pinhassi *et al.*, *Appl. Environ. Microbiol.* **71**, 7650 (2005).

42. I thank G. Wolfe for comments on the manuscript. This work was supported by NSF.

10.1126/science.1153527

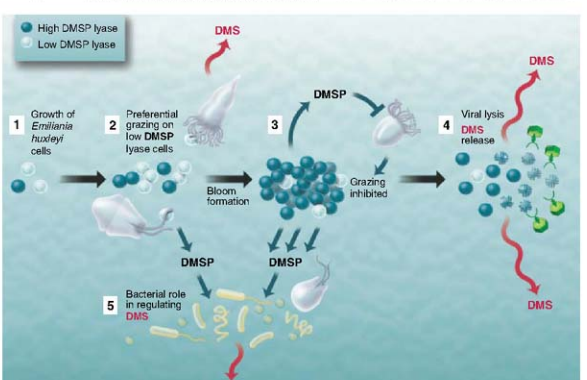


Fig. 2. Microbial community processes potentially regulating production of the climate-active gas DMS. (**1**) Environmental conditions promote growth of DMSP-containing *Emiliania huxleyi* cells. (**2**) Protist grazers respond by increasing abundance and grazing activity; grazing results in the release of DMS (*39*). Grazing deterrence by *E. huxleyi* cells with high DMSP lyase activity (shaded cells) allows an *E. huxleyi* bloom to form (*22*) and limits DMS output. (**3**) High concentrations of *E. huxleyi* cells are associated with high levels of DMSP release (*40*); dissolved DMSP inhibits protist grazing (*25*), promoting further bloom formation. (**4**) Viral lysis of cells in an aging bloom promotes DMS release (*24*). (**5**) DMSP is metabolized by bacteria including *Roseobacter*; propensity for the bacterial community to produce DMS versus demethylate DMSP is determined by community composition (*41*), shaped in part by differential mortality (e.g., due to nanoflagellate grazing selectivity) and competition for resources.

Extending the Sub–Sea-Floor Biosphere

Erwan G. Roussel,[1] Marie-Anne Cambon Bonavita,[1] Joël Querellou,[1] Barry A. Cragg,[2] Gordon Webster,[2] Daniel Prieur,[1] R. John Parkes[2]*

An extensive, global, sub–sea-floor biosphere has recently been documented (*1*), with the deepest sedimentary prokaryotes so far confirmed at 842 m depth, ~55°C, and 3.5 million years (My) old (*2*). It has been suggested that the sub–sea-floor biosphere may contain two-thirds of Earth's total prokaryotic biomass (*3*), but this extrapolation requires analysis of prokaryotic populations at greater depths. Here, we provide evidence for living prokaryotic cells in 1626 mbsf (meters below the sea floor) sediments that are 111 My old and at 60° to 100°C.

Nine deep sediment samples (*4*) (860 to 1626 mbsf) from the Newfoundland Margin [Ocean Drilling Program (ODP) leg 210, site 1276] were analyzed (*5*). These age from 46 to 111 My and are mainly hemipelagic mudrocks with interbedded gravity-flow deposits, formed during the opening of the northern Atlantic Ocean (*4*). An igneous sill at ~1613 mbsf (~10-m thick) is a trap for gases (methane, Fig. 1) and fluids. On the basis of the local thermal gradient and the measured thermal conductivity (*4*), the temperature of the deepest sample is calculated to be between 60° and 100°C. The upper temperature is close to the highest temperature for prokaryotic growth (~113°C) and where thermogenic alteration of organic matter can occur. The occurrence of thermogenic reactions and high temperatures (~80° to >100°C) in the deeper part of the core is supported by the presence of high methane concentrations and higher molecular weight hydrocarbons (*4*).

Intact prokaryotic cells were detected by microscopy in all samples, and many were dividing cells. Depth profiles of cell numbers fluctuate around 1.5×10^6 cells ml^{-1} (Fig. 1), consistent with the general depth distribution of prokaryotic cells from other sub–sea-floor sediments (*1*). Overall, cell numbers were significantly correlated ($P < 0.01$) with the organic matter hydrogen index (HI), which reflects organic matter reactivity for microorganisms (*6*) and implies that some of the cells were metabolically active. This is consistent with the presence of dividing cells (Fig. 1) and the detection of live cells using Live/Dead (Molecular Probes Incorporated, Eugene, OR) staining (~60% viable cells). Percentages of dividing cells of the total count are in the range of previous results (0 to 4.8%) (*1*), except for higher values in the deepest sample (1626 mbsf, 11.8%). This increase in the proportion of dividing cells is probably related to the dramatic increase in concentrations of methane and higher hydrocarbons together with the generally elevated organic carbon concentrations, both providing potential energy sources (Fig. 1) (*4*).

Successful archaeal 16*S* ribosomal RNA (rRNA) gene amplification also strongly suggests that the cells are from living prokaryotes because preservation of extracellular DNA in 46 to 111 My old, hot (60° to 100°C) sediments is unlikely. In addition, some detected sequences are related to thermophiles and/or hyperthermophiles, which matches in situ temperatures. The resulting clone libraries show a low diversity of Archaea (Fig. 1), with thermophilic *Pyrococcus* dominating the 958-m depth, and then, as soon as methane increases above background concentrations, potential anaerobic methane-oxidizing (ANME) sequences became dominant (fig. S1). Surprisingly, there were no ANME sequences in the deepest sample (1626 mbsf) despite the highest methane concentrations; however, at 60° to 100°C this may be above the upper temperature limit for ANME prokaryotes (*7*). In addition, the *Pyrococcus* and *Thermococcus* sequences (fig. S1) in the deepest sample may belong to high-temperature Archaea that are able to use the thermogenic higher hydrocarbons that accumulate below the sill.

These data provide direct evidence that significant prokaryotic populations are present in marine sediments at depths greater than a kilometer and as old as 111 My. This study also suggest that Archaea capable of anaerobic oxidation of methane and novel members of the high-temperature Thermococcales (*Pyrococcus* and *Thermococcus*) can dominate deep and hot sediments where there are thermogenic energy sources.

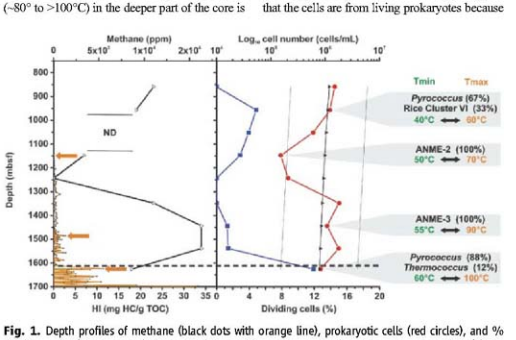


Fig. 1. Depth profiles of methane (black dots with orange line), prokaryotic cells (red circles), and % dividing cells (blue squares) determined by acridine orange staining. Regression equation (*1*) for prokaryotic cells in marine sediments (solid triangles), prediction limits (...). Orange arrows show local increases in methane. HI (open triangles) measured as mg of hydrocarbon (HC) per g of total organic carbon (TOC). ND, not determined. Dominant archaeal 16*S* rRNA gene sequences and in situ temperature range are on the right at the depths obtained. The diabase sill is shown as a bold horizontal dashed line.

References and Notes

1. R. J. Parkes *et al.*, *Nature* **371**, 410 (1994).
2. P. Wellsbury, I. Mather, R. J. Parkes, *FEMS Microbiol. Ecol.* **42**, 59 (2002).
3. W. B. Whitman, D. C. Coleman, W. J. Wiebe, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6578 (1998).
4. B. E. Tucholke *et al.*, *Drilling the Newfoundland Half of the Newfoundland-Iberia Transect: The First Conjugate Margin Drilling in a Nonvolcanic Rift*, vol. 210 of *Proceedings of the Ocean Drilling Program, Initial Reports* (Ocean Drilling Program, Texas A&M University, College Station, TX, 2004).
5. Materials and methods are available on *Science* Online.
6. R. J. Parkes *et al.*, *Org. Geochem.* **38**, 845 (2007).
7. J. Kallmeyer, A. Boetius, *Appl. Environ. Microbiol.* **70**, 1231 (2004).
8. We thank members of the Leg 210 cruise, especially J. C. Sibuet, for obtaining samples and K. E. Louden for helpful discussions and providing the thermal gradient data. This study used data provided by the ODP (www-odp.tamu.edu) and was partly funded from Natural Environment Research Council, UK. GenBank accession numbers are AM418589 to AM418604.

Supporting Online Material

www.sciencemag.org/cgi/content/full/320/5879/1046/DC1
Materials and Methods
Fig. S1

21 December 2007; accepted 10 March 2008
10.1126/science.1154545

[1]Laboratoire de Microbiologie des Environnements Extrêmes, UMR 6197, Université de Bretagne Occidentale, Ifremer, Centre de Brest, BP70, 29280 Plouzané, France. [2]School of Earth and Ocean Sciences, Cardiff University, Main Building, Park Place, Cardiff CF10 3YE, UK.

*To whom correspondence should be addressed. E-mail: J.Parkes@earth.cf.ac.uk

Virus Population Dynamics and Acquired Virus Resistance in Natural Microbial Communities

Anders F. Andersson* and Jillian F. Banfield†

Viruses shape microbial community structure and function by altering the fitness of their hosts and by promoting genetic exchange. The complexity of most natural ecosystems has precluded detailed studies of virus-host interactions. We reconstructed virus and host bacterial and archaeal genome sequences from community genomic data from two natural acidophilic biofilms. Viruses were matched to their hosts by analyzing spacer sequences that occur among clustered regularly interspaced short palindromic repeats (CRISPRs) that are a hallmark of virus resistance. Virus population genomic analyses provided evidence that extensive recombination shuffles sequence motifs sufficiently to evade CRISPR spacers. Only the most recently acquired spacers match coexisting viruses, which suggests that community stability is achieved by rapid but compensatory shifts in host resistance levels and virus population structure.

Viruses are arguably the most abundant and diverse components of natural environments. They can potentially alter the makeup, and thus the functioning, of microbial communities (*1*–*3*) and excise and transfer host DNA, facilitating genetic exchange and driving host evolution [e.g., (*4*, *5*)]. An important recent advance has been the recognition of a virus resistance system in bacteria and archaea that is based on genetically encoded spacers within the clustered regularly interspaced short palindromic repeats (CRISPR) loci (*6*–*8*). Experiments using *Streptococcus thermophilus* strains and their viruses established that CRISPR spacers with sequence identity to viral genomes confer specific resistance to viruses. Resistance is lost if there is even a single nucleotide mismatch between spacer and virus genome (*9*).

Although a suite of CRISPR-associated (Cas) proteins is implicated in microbial resistance (*10*, *11*), the details of the CRISPR-based mechanism and the dynamics of interactions involving viruses and the host CRISPR loci are not yet known. Community genomic analyses can capture roles for population heterogeneity in virus-microbe interactions in natural systems that are not apparent in pure culture experiments. Previously, we analyzed two closely related bacterial *Leptospirillum* populations and found patterns of variation consistent with very rapid evolution of the CRISPR locus (*12*). Given their likely viral origin (*6*–*9*), we hypothesized that spacers of CRISPR loci recovered from microbial communities could be used to fish out viral sequences from among the numerous, small, and otherwise unassigned community genomic fragments and to link viruses to their coexisting host bacteria and archaea (fig. S1A). Sufficient viral sequence was recovered to enable extensive reconstruction of multiple virus genomes from natural microbial communities.

DNA was extracted from two biofilms collected within the Richmond Mine, Redding, CA (fig. S1B). The subaerial UBA biofilm, dominated by bacterial *Leptospirillum* groups II (*13*) and III, was growing in 39°C pH 1.1 acid mine drainage (AMD) and was collected in June 2005. The floating UBA BS biofilm, growing in 39°C, pH 1.5 AMD, and collected in November 2005, contained abundant E-, G-, and A-plasma (*14*) and the more distantly related I-plasma, all uncultivated members of the Thermoplasmatales lineage of Euryarchaea. The sample also contained *Leptospirillum* groups II and III and other novel archaea (*14*). Approximately 100 Mb of genomic sequence was obtained from each of the small insert libraries constructed from the UBA and UBA BS samples (*15*). Essentially complete genomes of *Leptospirillum* groups II (*13*) and III were reconstructed from the UBA sample. In addition, near-complete genomes of I-plasma (~20-fold coverage) and E-plasma (~10-fold coverage) and lower coverage but near-complete genomes of G-plasma and A-plasma were reconstructed from sequencing reads derived from both samples (*15*).

In the combined community genomics data set, we identified 476 reads encoding 37 different CRISPR repeat sequences (fig S1A and table S1). A subset of the sequences were mapped to CRISPR loci of *Leptospirillum* group II (*16*), *Leptospirillum* group III, I-plasma, E-plasma, A-plasma, and G-plasma genomes and bacterial plasmids (*15*). From the CRISPR reads, 6044 spacer sequences, representing 2348 unique sequences (28 to 54 nucleotides in length) were extracted (table S2). As initial confirmation that reads exactly matching CRISPR spacers (and

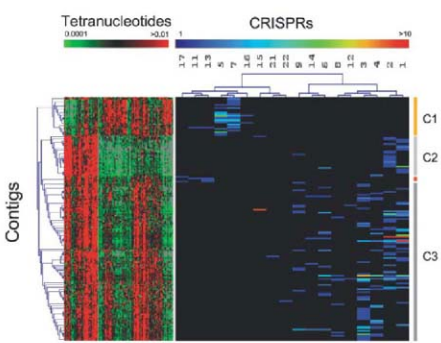


Fig. 1. Virus-host associations in AMD biofilms. Putative viral (SNC) contigs were clustered based on tetranucleotide frequencies (left panel), and CRISPRs were clustered based on patterns of SNC contig matching (right panel) (*15*). Columns in the left panel represent tetranucleotides (reverse complementary pairs are combined); colors indicate frequencies (gray indicates absence). Columns in the right panel represent CRISPRs; colors indicate number of distinct spacer sequences matching SNC contigs. The majority of Cluster 1 (C1) contigs belong to the AMDV1 population, Cluster 2 (C2) to AMDV2, and Cluster 3 (C3) to AMDV3, AMDV4, and AMDV5 (see table S3 for details).

Departments of Earth and Planetary Science and Environmental Science, Policy, and Management, University of California, Berkeley, CA 94720, USA.

*Present address: Department of Ecology and Evolution, Uppsala University, and Swedish Institute for Infectious Disease Control, SE-17182 Solna, Sweden.

†To whom correspondence should be addressed. E-mail: jbanfield@berkeley.edu

lacking repeats) derive from coexisting viruses, all spacer-containing non-CRISPR (SNC) reads were compared with sequences in public databases. The 911 SNC reads were significantly enriched in genes with matches to viral proteins, compared with random reads (Fisher's exact test, $P < 10^{-16}$) (fig. S2), and were also enriched in genes encoding proteins with typical viral functions (fig. S3). Thus, the majority of SNC reads derived from viruses, although some corresponded to other mobile genetic elements, such as plasmids and transposons.

Exact matches to spacer sequences were also found on assembled DNA fragments (contigs) up to 15 kb in length. SNC contigs were typically linked by clone mate pairs to other SNC contigs, but only rarely to archaeal or bacterial genome fragments (supporting online text). On the basis of mate-pair linkage and blocks of similar sequence, SNC contigs were condensed by manual assembly (*15*) to form larger fragments that we inferred represented partial and possibly complete viral genomes.

On the basis of tetranucleotide frequencies (*17*), SNC contigs were clustered into three major groups (Fig. 1). The first cluster was formed from contigs that match CRISPR spacers of the bacterial *Leptospirillum* groups II and III. The second and third clusters were targeted by CRISPR spacers of archaea and of plasmids (supporting online text).

Many of the SNC contigs of the first tetranucleotide cluster (Fig. 1) were linked by mate pairs into subclusters (table S3). One such subcluster (AMDV1) was targeted by *Leptospirillum* group II and III spacers. AMDV1 is similar to prophages integrated into the genomes of *Gluconobacter oxydans* 621H (*18*) and *Acidiphilum cryptum* JF-5 (fig. S4 and supporting online text).

The second tetranucleotide cluster (Fig. 1) includes a deeply sampled SNC contig (AMDV2) (fig. S6 and supporting online text) and smaller contigs linked to it by mate pairs and overlapping sequence (fig. S5 and table S3). AMDV2 is targeted by 33 spacers derived only from the two E-plasma CRISPR loci. The composite sequence of the discrete, linear 10-kb viral genome has a GC content of 23.7% and inverted 160 base pair (bp) repeats on each end, and encodes 17 (putative) genes, including a type B DNA polymerase, all on the same strand (Fig. 2A).

The third tetranucleotide cluster (Fig. 1) comprises several subclusters of mate-pair-linked and partially overlapping contigs (fig. S5 and table S3). One subcluster, AMDV3, includes ~18 kb of composite genome sequence plus many small strain variant contigs (fig. S7). Given targeting of AMDV3 by the CRISPR loci of E-, G-, and A-plasma [88 to 92% 16S ribosomal RNA (rRNA) gene sequence identity], we infer that this virus population has a relatively broad host range.

AMDV4 (~56 kb composite genome) shares some sequence similarity with AMDV3 (fig. S5) and groups with it in tetranucleotide cluster 3. E-plasma is the likely host for this virus (table S3 and supporting online text). Although the E-plasma type #1 CRISPR locus targets AMDV4, more of its spacers target AMDV2. The interspersing of spacers matching to AMDV2 and AMDV4 on single CRISPR reads suggests that E-plasma is exposed to both of these narrow host range viruses simultaneously.

Only one virus, AMDV5, targets I-plasma, and it is only detected in the UBA BS sample that contains this archaeon. I-plasma shares only 80% 16S rRNA gene sequence identity with the closest of the other Thermoplasmatales archaea and has no spacers that would confer immunity to other Thermoplasmatales viruses (table S3). All 10 I-plasma spacers that match SNC reads colocalize to the two AMDV5 fragments (~4 kb). Interestingly, three AMDV5 proteins have closest matches in eukaryal viruses (supporting online text).

CRISPR loci 11, 13, 17, and 19 are associated with plasmid-like populations (table S3 and supporting online text). These complex populations may be responsible for CRISPR lateral transfer (*19*).

Up to 40% of spacers in a single CRISPR locus matched virus sequences, the highest degree of correspondence being for the AMDV2 population. Given that each CRISPR read represents an individual cell, it is evident that most microbial cells target several different virus populations (e.g., some E-plasma cells target specific AMDV2, 3, and 4 variants). By mapping spacers onto reads coassembled into AMDV2 and AMDV3, we found that some spacers match the dominant sequence types, yet many match sequences characteristic of only one or a small number of genotypes within each population (e.g., fig. S8).

Comparative analyses of isolated bacteriophages reveal genome mosaicism (*20*–*24*). Less

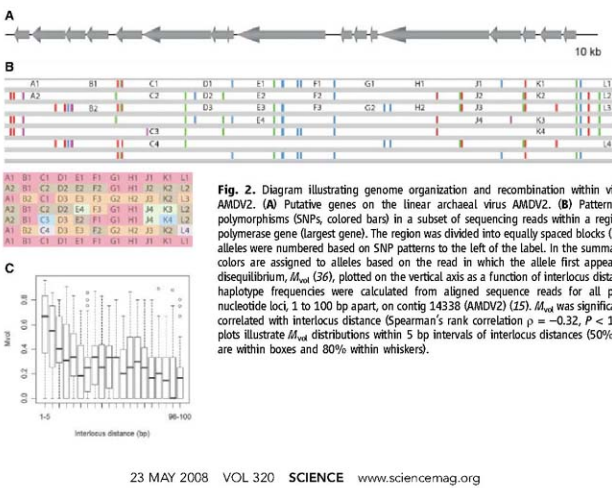


Fig. 2. Diagram illustrating genome organization and recombination within virus population AMDV2. **(A)** Putative genes on the linear archaeal virus AMDV2. **(B)** Pattern of nucleotide polymorphisms (SNPs, colored bars) in a subset of sequencing reads within a region of the DNA polymerase gene (largest gene). The region was divided into equally spaced blocks (A to L), and the alleles were numbered based on SNP patterns to the left of the label. In the summary table below, colors are assigned to alleles based on the read in which the allele first appears. **(C)** Linkage disequilibrium, M_{vol} (*36*), plotted on the vertical axis as a function of interlocus distance. Allele and haplotype frequencies were calculated from aligned sequence reads for all pairs of single-nucleotide loci, 1 to 100 bp apart, on contig 14338 (AMDV2) (*15*). M_{vol} was significantly negatively correlated with interlocus distance (Spearman's rank correlation $\rho = -0.32$, $P < 10^{-16}$). The box plots illustrate M_{vol} distributions within 5 bp intervals of interlocus distances (50% of data points are within boxes and 80% within whiskers).

is known about natural population heterogeneity or recombination in archaeal viruses [but see e.g., (*25*–*29*)]. The AMDV2 population displays a high level of nucleotide variation (~94% average similarity of aligned reads to each other). Combinatorial mixtures of small sequence motifs (Fig. 2B and fig. S9) suggest that the population has been shaped by extensive homologous recombination. We measured linkage disequilibrium as a function of interlocus distance and found a significant decline in linkage with distance (Spearman's correlation, $P < 10^{-16}$) (Fig. 2C). Linkage becomes independent of distance at loci separations of >25 nucleotides (values remain close to 0.1). Thus, recombination has scrambled virus sequences so much that blocks shared by different individuals are often no more than 25 nucleotides in length, in agreement with the size of sequence motifs apparent in Fig. 2B. The persistence of some linkage over longer distances (Fig. 2C) is due to a low abundance of highly similar sequences. Recombination creates new virus sequences that are able to confound the function of 28 to 54 nucleotide CRISPR spacers, allowing the virus to evade the CRISPR-based host defense system (fig. S8). Compared with mutation, recombination generates new DNA signatures with less risk of altering protein function and limits purging to sequence motifs recognized by spacers rather than entire viral genotypes, preserving viral population diversity.

Within the AMDV3 and AMDV4 populations, genotypes share larger sequence motifs than AMDV2 (Fig. 3). In AMDV3, substitution of highly divergent sequence blocks within some variants of a large membrane or capsid-like protein gene (Fig. 3 and fig. S7) may alter the host range, allowing infection of the A-, E-, and G-plasma lineages, analogous to specificity conferred by recombination in bacteriophage tail fiber proteins (*30*).

The set of CRISPR spacers in each cell within the host populations differs, and only a few CRISPR spacers are shared widely (table S4). The exception is I-plasma, where the first ~800 bp of the CRISPR locus is clonal (i.e., each cell has the same spacer in the same position as every other cell for the first 12 of, on average, ~15 spacers). The CRISPR locus diversifies quickly toward the *cas* genes; the terminal region provided 36 single-copy spacers. The colocation of the switch to nonclonal spacers and to spacers that match AMDV5 (supporting online text) suggests that the recent appearance of AMDV5 caused a selective sweep. The fact that mostly rare and recently incorporated (close to the *cas* genes) spacers (*7*, *9*, *12*, *31*) match the I-plasma virus (11 of 13 are in single copy) also applies to other archaea (table S4).

The complexity of most previously studied systems has limited resolution of patterns of viral and host distribution over space and time. In the current study, most virus populations (AMDV2–4) and their hosts were present in the two samples collected about 5 months apart. However, the relative abundances of AMDV2 and AMDV4 populations, inferred to primarily target the same host (E-plasma), differed between samples (supporting online text). Interestingly, in the June sample, all of the matching spacers were found on the type 1 CRISPR locus, whereas the majority in the November sample were carried on the type 2 locus, which had expanded considerably (197 versus 37 distinct spacer sequences). Moreover, there was only one matching spacer common to the E-plasma populations sampled at the two time points. This suggests rapid evolution of the CRISPR loci and potential modulation of resistance levels on the time scale of months.

CRISPR loci cannot grow unchecked (*12*). Consequently, a host may be exposed to a virus for which it has no spacer-based immunity as the result of spacer loss. Alternatively, this may occur as a result of migration of a new virus type into the community (*32*) or evolution of a virus to predate a new host. For infections that lower the fitness of the host, the first cells either to acquire an effective CRISPR locus by lateral transfer (e.g., from the plasmid pool) or a spacer matching the new virus, would thrive relative to other individuals in its population. Proliferation of these cells will spread effective spacers in the population, generating a pattern comparable to that observed in the I-plasma CRISPR locus. Increase in population-level immunity by spacer acquisition (*33*) will be countered by virus mutation (*33*) and recombination, and possibly other mechanisms of genome evolution (*34*). Resistance may increase over time to the point that the virus population declines, or a virus may occasionally become so virulent that it causes a crash in the host population, as predicted by the "kill the winner" model (*35*). Alternatively, if CRISPR and viral diversification remain in balance, a relatively stable virus and host community may result.

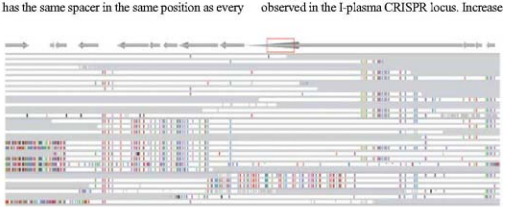


Fig. 3. Diagram illustrating heterogeneity within a genomic region of the archaeal virus population, AMDV3. Genes (row of arrows) and a subset of the assembled sequencing reads on contig 15322 (bars are reads, thin white lines link mate pairs) are shown. The red box identifies the region of the gene magnified below, where colored bars indicate single-nucleotide polymorphisms (SNPs). Linkage patterns among defined SNP patterns are consistent with extensive recombination.

References and Notes

1. J. A. Fuhrman, *Nature* **399**, 541 (1999).
2. G. Bratbak, M. Heldal, S. Norland, T. F. Thingstad, *Appl. Environ. Microbiol.* **56**, 1400 (1990).
3. M. Breitbart, F. Rohwer, *Trends Microbiol.* **13**, 278 (2005).
4. C. Canchaya, G. Fournous, S. Chibani-Chennoufi, M. L. Dillmann, H. Brussow, *Curr. Opin. Microbiol.* **6**, 417 (2003).
5. H. Ochman, J. G. Lawrence, E. A. Groisman, *Nature* **405**, 299 (2000).
6. F. J. Mojica, C. Diez-Villasenor, J. Garcia-Martinez, E. Soria, *J. Mol. Evol.* **60**, 174 (2005).
7. C. Pourcel, G. Salvignol, G. Vergnaud, *Microbiology* **151**, 653 (2005).
8. A. Bolotin, B. Quinquis, A. Sorokin, S. D. Ehrlich, *Microbiology* **151**, 2551 (2005).
9. R. Barrangou *et al.*, *Science* **315**, 1709 (2007).
10. D. H. Haft, J. Selengut, E. F. Mongodin, K. E. Nelson, *PLOS Comput. Biol.* **1**, e60 (2005).
11. K. S. Makarova, N. V. Grishin, S. A. Shabalina, Y. I. Wolf, E. V. Koonin, *Biol. Direct* **1**, 7 (2006).
12. G. W. Tyson, J. F. Banfield, *Environ. Microbiol.* **10**, 200 (2007).
13. I. Lo *et al.*, *Nature* **446**, 537 (2007).
14. G. K. Druschel, B. J. Baker, T. H. Gihring, J. F. Banfield, *Geochem. Trans.* **5**, 13 (2004).
15. Materials and methods are available as supporting material on *Science* Online.
16. G. W. Tyson *et al.*, *Nature* **428**, 37 (2004).
17. D. T. Pride, T. M. Wassenaar, C. Ghose, M. J. Blaser, *BMC Genomics* **7**, 8 (2006).
18. C. Prust *et al.*, *Nat. Biotechnol.* **23**, 195 (2005).
19. J. S. Godde, A. Bickerton, *J. Mol. Evol.* **62**, 718 (2006).
20. D. Botstein, *Ann. N. Y. Acad. Sci.* **354**, 484 (1980).
21. A. S. Nilsson, E. Haggard-Ljungquist, *Mol. Phylogenet. Evol.* **21**, 259 (2001).
22. R. J. Juhala *et al.*, *J. Mol. Biol.* **299**, 27 (2000).
23. M. L. Pedulla *et al.*, *Cell* **113**, 171 (2003).
24. H. Brussow, F. Desiere, *Mol. Microbiol.* **39**, 213 (2001).
25. X. Peng *et al.*, *Virology* **291**, 226 (2001).
26. B. Wiedenheft *et al.*, *J. Virol.* **78**, 1954 (2004).
27. G. Vestergaard *et al.*, *J. Virol.* **82**, 371 (2007).
28. S. L. Tang, S. Nuttall, M. Dyall-Smith, *J. Bacteriol.* **186**, 2810 (2004).
29. N. Rossler, R. Klein, H. Scholz, A. Witte, *Mol. Microbiol.* **52**, 413 (2004).
30. F. Tetart, C. Desplats, H. M. Krisch, *J. Mol. Biol.* **282**, 543 (1998).
31. R. K. Lillestol, P. Redder, R. A. Garrett, K. Brugger, *Archaea* **2**, 59 (2006).
32. J. C. Snyder *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 19102 (2007).
33. H. Deveau *et al.*, *J. Bacteriol.* **190**, 1390 (2008).
34. X. Peng, A. Kessler, H. Phan, R. A. Garrett, D. Prangishvili, *Mol. Microbiol.* **54**, 366 (2004).
35. T. F. Thingstad, *Limnol. Oceanogr.* **45**, 1320 (2000).
36. Y. Chen, C. H. Lin, C. Sabatti, *BMC Genet.* **7**, 54 (2006).

37. We thank M. Breitbart, P. Wilmes, V. Denef, P. Unneberg, B. Weidenheft, S. Simmons, B. Baker and C. Sun for comments and suggestions. This research was supported

by the U.S. Department of Energy Genomics GTL program. A.A. was supported by a grant from the Swedish Research Council. We thank T. Arman (Iron Mountain Mines), R. Carver, and R. Sugarek for site access and other assistance. DNA sequencing was performed at the Department of Energy's Joint Genome Institute. This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the project accession ABOZ00000000. The version described in this paper is the first version, ABOZ01000000.

Supporting Online Material
www.sciencemag.org/cgi/content/full/320/5879/1047/DC1
Materials and Methods
SOM Text
Figs. S1 to S9
Tables S1 to S4

5 March 2008; accepted 23 April 2008
10.1126/science.1157358

Regulated Protein Denitrosylation by Cytosolic and Mitochondrial Thioredoxins

Moran Benhar,[1] Michael T. Forrester,[2] Douglas T. Hess,[1] Jonathan S. Stamler[1,2]*

Nitric oxide acts substantially in cellular signal transduction through stimulus-coupled *S*-nitrosylation of cysteine residues. The mechanisms that might subserve protein denitrosylation in cellular signaling remain uncharacterized. Our search for denitrosylase activities focused on caspase-3, an exemplar of stimulus-dependent denitrosylation, and identified thioredoxin and thioredoxin reductase in a biochemical screen. In resting human lymphocytes, thioredoxin-1 actively denitrosylated cytosolic caspase-3 and thereby maintained a low steady-state amount of *S*-nitrosylation. Upon stimulation of Fas, thioredoxin-2 mediated denitrosylation of mitochondria-associated caspase-3, a process required for caspase-3 activation, and promoted apoptosis. Inhibition of thioredoxin-thioredoxin reductases enabled identification of additional substrates subject to endogenous *S*-nitrosylation. Thus, specific enzymatic mechanisms may regulate basal and stimulus-induced denitrosylation in mammalian cells.

Cellular functions of nitric oxide (NO) are carried out in part through *S*-nitrosylation of Cys residues within a broad functional spectrum of proteins (*1*). Like protein phosphorylation, *S*-nitrosylation thus mediates or modulates transduction of myriad cellular signals. Although the NO synthases (NOSs) and *S*-nitrosoglutathione reductase (GSNOR) govern *S*-nitrosylation by NO or *S*-nitrosothiols (SNOs) (*1*), little is known about the nature of or even necessity for enzymatic mechanisms that may directly remove NO groups from Cys thiols. Indeed, stimulus-coupled denitrosylation has been shown to precisely modulate the function of proteins that are constitutively *S*-nitrosylated (*2*, *3*), but the cellular effectors of denitrosylation remain undetermined.

Some caspases, members of a family of proteases that mediate apoptosis, are subject to inhibitory *S*-nitrosylation, and stimulation of death receptors of the tumor necrosis factor superfamily results in caspase denitrosylation (*2*, *4*, *5*). In particular, a subpopulation of caspase-3 (a major executioner caspase) that is associated with mitochondria is constitutively *S*-nitrosylated at the active site Cys, and engagement of the Fas receptor promotes denitrosylation (*2*, *6*). However, the role of denitrosylation in transduction of the apoptotic signal has not been fully elucidated, and the molecular mechanism of denitrosylation is unknown.

Identification of thioredoxin as a caspase-3 denitrosylase. To assess protein denitrosylation, we developed an assay based on the reactivation of caspase that results from denitrosylation of the active site Cys (*7*) (fig. S1A). Recombinant caspase-3 was *S*-nitrosylated (SNO–caspase-3), and the inhibited enzyme was immobilized on nickel-coated plates. Bound SNO–caspase-3 was incubated with cellular extracts, and caspase-3 activity was then measured with a fluorogenic

[1]Department of Medicine, Duke University Medical Center, Durham, NC 27710, USA. [2]Department of Biochemistry, Duke University Medical Center, Durham, NC 27710, USA.

*To whom correspondence should be addressed. E-mail: stam1001@mc.duke.edu

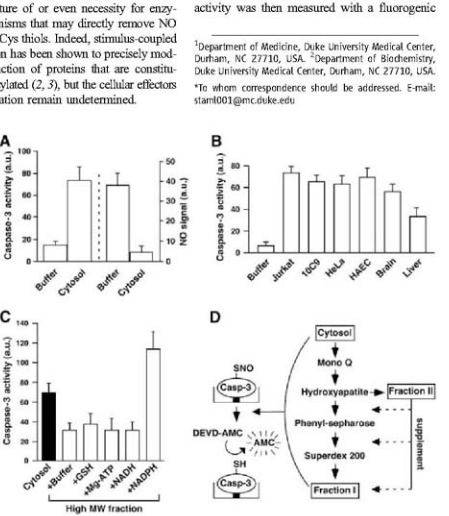


Fig. 1. Characterization of an enzymatic activity that denitrosylates SNO–caspase-3. Data in (**A**) to (**C**) are presented as mean ± SEM; $n \geq 3$. (**A**) Reactivation of SNO–caspase-3 protease activity by cell cytosol. Immobilized SNO–caspase-3 (~100 nM) was incubated for 30 min with a cytosolic extract prepared from Jurkat cells. (left) Caspase activity was determined by using Z-DEVD-7-amino-4-methylcoumarin (DEVD-AMC). (Right) SNO content was assayed by using Hg-coupled photolysis-chemiluminescence. a.u., arbitrary unit. (**B**) Caspase activity after 30-min incubation of SNO–caspase-3 with cytosolic fractions (100 μg of protein) prepared from human cells or rat tissues. HAEC, primary human aortic endothelial cells. (**C**) SNO–caspase-3 was incubated with the cytosolic fraction from Jurkat cells or with cytosolic fractions after size-exclusion chromatography through Sephadex G-25 [high molecular weight (MW) fraction] and supplemented with GSH (0.5 mM), ATP (10 μM), NADH (10 μM), or NADPH (10 μM). (**D**) The procedure used to partially purify a SNO–caspase-3 denitrosylase (fig. S1, A and E, and table S1).

substrate. Addition of Jurkat cell cytosolic extract rapidly restored the activity of caspase-3 (Fig. 1A). Loss of the SNO moiety was verified by Hg-coupled photolysis-chemiluminescence, which measures the amount of NO displaced from Cys thiol by ultraviolet (UV) light (*2*) (Fig. 1A). This cytosolic denitrosylating activity was detected in numerous human cell types and in rat tissues (Fig. 1B). A subcellular fraction enriched in mitochondria also displayed denitrosylating activity (fig. S1B).

Cytosolic denitrosylating activity was present in a fraction enriched for large molecular size (>10 kD) and was diminished after exposure to heat or trypsin (fig. S1C). Activity was also reduced after fractionation by size-exclusion chromatography (Sephadex G-25), indicating a possible requirement for a small cofactor, and activity was restored and potentiated by reduced nicotinamide adenine dinucleotide phosphate (NADPH) but not nicotinamide dinucleotide (NADH), glutathione (GSH), or adenosine triphosphate (ATP) (Fig. 1C). Thus, the cytosolic denitrosylating activity exhibits properties of an NADPH-dependent oxidoreductase.

We derived, by a four-step chromatographic purification from Jurkat cells, a highly active fraction (designated fraction I) (Fig. 1D and table S1), whose activity was dependent on a second fraction added in limiting amounts (1:10) (designated fraction II) (Fig. 1D and fig. S1D). Fraction I contained eight proteins (fig. S1E), which were identified by matrix-assisted laser

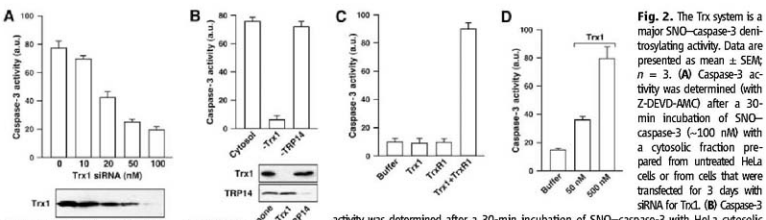


Fig. 2. The Trx system is a major SNO–caspase-3 denitrosylating activity. Data are presented as mean ± SEM; $n = 3$. **(A)** Caspase-3 activity was determined (with Z-DEVD-AMC) after a 30-min incubation of SNO–caspase-3 (~100 nM) with a cytosolic fraction prepared from untreated HeLa cells or from cells that were transfected for 3 days with siRNA for Trx1. **(B)** Caspase-3 activity was determined after a 30-min incubation of SNO–caspase-3 with HeLa cytosolic extract or cytosol that had been depleted of Trx1 or TRP14 by using specific antibodies against Trx1 or TRP14. **(C)** Caspase-3 activity was determined after a 30-min incubation of SNO–caspase-3 (~100 nM) with NADPH (100 μM) and recombinant human Trx1 (10 nM) and/or recombinant rat TrxR1 (10 nM). **(D)** Caspase-3 activity after a 30-min incubation with recombinant *Escherichia coli* Trx1.

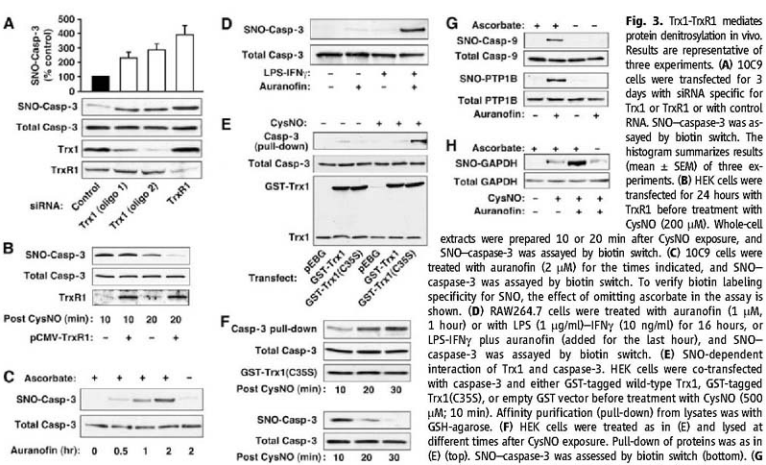


Fig. 3. Trx1-TrxR1 mediates protein denitrosylation in vivo. Results are representative of three experiments. **(A)** 10C9 cells were transfected for 3 days with siRNA specific for Trx1 or TrxR1 or with control RNA. SNO–caspase-3 was assayed by biotin switch. The histogram summarizes results (mean ± SEM) of three experiments. **(B)** HEK cells were transfected for 24 hours with TrxR1 before treatment with CysNO (200 μM). Whole-cell extracts were prepared 10 or 20 min after CysNO exposure, and SNO–caspase-3 was assayed by biotin switch. **(C)** 10C9 cells were treated with auranofin (2 μM) for the times indicated, and SNO–caspase-3 was assayed by biotin switch. To verify biotin labeling specificity for SNO, the effect of omitting ascorbate in the assay is shown. **(D)** RAW264.7 cells were treated with auranofin (1 μM, 1 hour) or with LPS (1 μg/ml)–IFNγ (10 ng/ml) for 16 hours, or LPS-IFNγ plus auranofin (added for the last hour), and SNO–caspase-3 was assayed by biotin switch. **(E)** SNO-dependent interaction of Trx1 and caspase-3. HEK cells were co-transfected with caspase-3 and either GST-tagged wild-type Trx1, GST-tagged Trx1(C35S), or empty GST vector before treatment with CysNO (500 μM; 10 min). Affinity purification (pull-down) from lysates was with GSH-agarose. **(F)** HEK cells were treated as in (E) and lysed at different times after CysNO exposure. Pull-down of proteins was as in (E) (top). SNO–caspase-3 was assessed by biotin switch (bottom). **(G** and **H)** 10C9 cells were treated with auranofin (2 μM) or vehicle [dimethyl sulfoxide (DMSO)] for 2 hours, and *S*-nitrosylation of endogenous caspase-9, PTP1B, and GAPDH was assessed by biotin switch. For GAPDH (H), CysNO (500 μM; 20 min) was added after auranofin treatment.

desorption ionization time-of-flight mass spectrometry (table S2). Of these, only thioredoxin-1 (Trx1) could be ascribed a redox-related function. Recombinant Trx reductase (TrxR) could substitute for fraction II, fully reconstituting the denitrosylating activity of fraction I (fig. S1D).

Depletion of Trx1 from HeLa cells with small interfering RNA (siRNA) correlated with a loss of SNO–caspase-3 denitrosylating activity in vitro (Fig. 2A and fig. S2A), and denitrosylating activity was restored by adding back recombinant Trx1 but not an active site mutant Trx [$Cys^{32} \rightarrow Ser^{32}$, Trx1(C32S)] (fig. S2A). In contrast, siRNA-mediated depletion of an additional member of the Trx family, Trx-related protein 14 (TRP14) (*8*), had no effect on denitrosylating activity (fig. S2B). Similarly, immunodepletion of Trx1 but not TRP14 abolished denitrosylating activity (Fig. 2B). A reconstituted Trx system [10 nM Trx and TrxR (Trx-TrxR) and including NADPH] efficiently denitrosylated an excess of SNO–caspase-3 (Fig. 2C). Denitrosylation by Trx1 in the absence of TrxR1 was ineffective but was restored when concentrations of Trx1, but not Trx1(C32S), approached or exceeded that of SNO–caspase-3 (Fig. 2D and fig. S2C), suggestive of single-turnover denitrosylation coupled to Trx1 oxidation.

Dynamic regulation of cellular protein *S*-nitrosylation by Trx-TrxR. Vicinal dithiols, as present in Trx (CXXC, where C represents Cys and X is another amino acid), exhibit a denitrosylating activity in vitro (*9*), and various redox enzymes, including dithiol proteins, can catalyze denitrosylation reactions ex vivo (*10*–*13*). Trx may also break down SNOs within cells exposed to high concentrations of a nitrosylating agent (*14*, *15*), as might occur during a nitrosative stress. However, in the context of signaling by endogenously derived NO, Trx has been proposed to promote *S*-nitrosylation of proteins (including caspase-3) (*16*, *17*). Thus, in sum, it is not known whether Trx might have a physiological role in mediating either basal or stimulus-induced denitrosylation in vivo.

The extent of *S*-nitrosylation reflects the equilibrium between addition and removal of NO groups. We used RNA interference to assess whether Trx governs basal *S*-nitrosylation of caspase-3 in human B lymphocytes (10C9 cells), where most caspase-3 is cytosolic and in the reduced form (i.e., only the minor population of caspase-3 associated with the mitochondria is constitutively *S*-nitrosylated) (*6*). Two nonoverlapping siRNAs efficiently decreased the amount of cellular Trx1, and amounts of endogenous SNO–caspase-3 [assessed by the biotin switch technique, in which *S*-nitrosylated cysteines are selectively biotinylated (*18*, *19*)] were increased in proportion to Trx1 depletion (Fig. 3A). siRNA-mediated depletion of iNOS or NOS inhibition with N^G-monomethyl-L-arginine (L-NMMA) prevented the increase in SNO–caspase-3 (fig. S3, A and B). Depletion of TrxR1 in 10C9 cells (Fig. 3A and fig. S3C) and in human embryonic kidney (HEK) cells stably expressing a second NOS isoform, nNOS (HEK-nNOS) (fig. S3D), also resulted in increased SNO–caspase-3, even more efficiently than did depletion of Trx1. Thus, TrxR1 activity may be rate-limiting for denitrosylation of cytosolic SNO–caspase-3 in vivo. Inhibition of Trx or TrxR might alter the cellular redox balance (e.g., decrease intracellular GSH levels) and thus affect the abundance of SNO-proteins independently of specific denitrosylating activity (*19*). However, knockdown of Trx1 or TrxR1 did not alter levels of acid soluble thiols (which consist primarily of GSH) (fig. S3E).

Equivalent amounts of SNO–caspase-3 accumulated in HEK cells overexpressing either TrxR1 or empty vector shortly after treatment with *S*-nitrosocysteine (CysNO), a cell-permeable *S*-nitrosylating compound (Fig. 3B; 10 min). However, after 20 min, the amount of SNO–caspase-3 was lower in TrxR1-transfected cells (Fig. 3B). Therefore, augmenting TrxR1 enhances SNO–caspase-3 denitrosylation.

We examined in intact cells the effects on denitrosylation of 1-chloro-2,4-dinitrobenzene (DNCB) or *S*-triethylphosphine-gold(I)-2,3,4,6-tetra-*O*-acetyl-1-thio-β-D-glucopyranoside (auranofin), two rapidly acting and potent but structurally dissimilar inhibitors of TrxR (*20*, *21*). Treatment of 10C9 cells with either DNCB or auranofin increased the amount of SNO–caspase-3 in cells exposed to CysNO (fig. S3F). Remarkably, in the absence of an exogenous NO donor, cells treated with auranofin also accumulated SNO–caspase-3 (Fig. 3C and fig. S3G). We validated the identification of endogenous SNO in assays with auranofin by showing the dependence of signals on ascorbate (*18*, *19*) as well as by the elimination of signals after UV irradiation, which liberates NO from SNO-proteins (*19*) (Fig. 3C and fig. S3H), and we observed that auranofin did not alter the abundance of acid-soluble thiols (fig. S3E). Thus, constitutive Trx1 activity apparently controls the steady-state amount of caspase-3 *S*-nitrosylation, a role similar to that of phosphatases in controlling basal phosphorylation.

To explore the role of Trx in reversing stimulus-coupled caspase-3 *S*-nitrosylation, we treated RAW264.7 macrophages with lipopolysaccharide (LPS) plus interferon-γ (IFNγ), which induces iNOS. Whereas LPS-IFNγ or auranofin alone weakly elicited *S*-nitrosylation of caspase-3, robust *S*-nitrosylation was seen when the two treatments were combined (Fig. 3D). Thus, *S*-nitrosylation of caspase-3 appears to reflect dynamic turnover of the NO group, mediated by the opposing activities of NOS and Trx-TrxR.

The reduction of substrate disulfides by Trx involves a CXXC motif (Cys^{32} and Cys^{35} in human Trx1) (*22*) (fig. S4A). A C35S mutant Trx can be used to trap substrates in a mixed disulfide (*23*). We developed an analogous SNO-based trapping strategy to examine the interaction of Trx with SNO–caspase-3. We transfected HEK cells with caspase-3 and Trx1(C35S) fused to glutathione *S*-transferase (GST). After treatment with CysNO, caspase-3 co-precipitated with Trx1(C35S) but not with wild-type Trx1 (Fig. 3E). The amount of caspase-3 trapped by Trx1(C35S) increased from 10 to 30 min after treatment with CysNO, during which time denitrosylation occurred (Fig. 3F). Additionally, in HEK cells transiently expressing iNOS, Trx(C35S) but not wild-type Trx formed mixed disulfides with multiple proteins in an NO-dependent manner (fig. S4D). Thus, Trx1 appears to interact with SNO–caspase-3 in vivo and to denitrosylate via a mixed-disulfide intermediate (fig. S4B); however, a mechanism involving transnitrosylation (*14*) is not excluded (fig. S4C).

Trx might serve as a denitrosylase for a broad spectrum of proteins. To identify endogenously *S*-nitrosylated substrates of Trx under basal conditions, we inhibited Trx in 10C9 cells with auranofin and identified caspase-9 and protein tyrosine phosphatase 1B (PTP1B) as substrates for Trx-mediated denitrosylation (Fig. 3G). The *S*-nitrosylated form of glyceraldehyde-3-phosphate dehydrogenase (GAPDH), whose proapoptotic function is induced by NO (*24*), was not detected in untreated cells and did not appreciably accumulate 2 hours after exposure to auranofin. However, treatment with CysNO, representing a nitrosative stress, resulted in accumulation of SNO-GAPDH, and the amounts of SNO-GAPDH were further increased markedly by auranofin (Fig. 3H). Similarly, inhibition of TrxR or Trx increased the abundance of additional SNO-proteins after exposure to CysNO (fig. S5, A and B) as well as slowed their denitrosylation (fig. S5, C and D). Further, in HEK-nNOS cells, depletion of TrxR1 with siRNA increased the amounts of whole-cell SNO (Fig. S5E). Collectively, these studies suggest that Trx-TrxR acts as a multisubstrate denitrosylase.

Regulation of caspase-3 denitrosylation and apoptosis by mitochondrial Trx-TrxR. We then asked whether the Trx system also mediates stimulus (Fas)-induced denitrosylation, first described for mitochondria-associated SNO–procaspase-3 in 10C9 cells (*2*, *6*). Both siRNA-mediated depletion of TrxR2, the isoform of TrxR that is localized selectively to mitochondria (*25*), and auranofin [which effectively inhibits TrxR2 (*26*)] abrogated Fas-dependent denitrosylation of mitochondrial caspase-3 (Fig. 4, A and B, and fig. S6A). Thus, Fas-induced denitrosylation of mitochondrial SNO–caspase-3 appears to be mediated by Trx2-TrxR2, a conclusion supported further by the finding that Fas stimulation resulted in an increase in mitochondrial Trx2 [152% ± 12% (mean ± SEM) relative to control, $n = 4$] (fig. S6B). These results suggest the possibility that Trx2-mediated denitrosylation [acting in concert with cleavage by initiator caspase(s)] may promote full activation of caspase-3 and thereby facilitate apoptosis.

We further examined this possibility by assessing the effects of mitochondrial TrxR2 knockdown or inhibition on two molecular events that characterize the execution phase of Fas- and

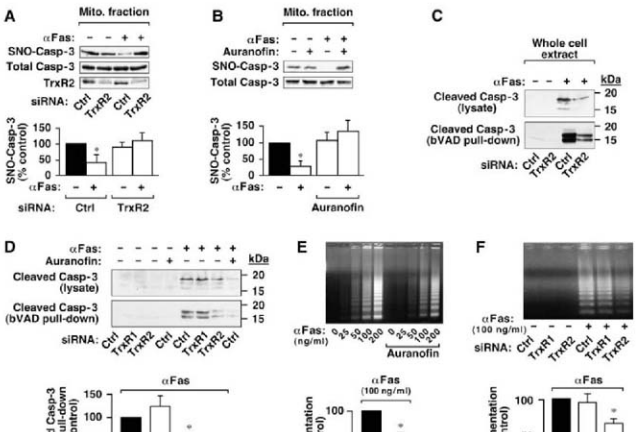


Fig. 4. The mitochondrial Trx system mediates Fas-induced denitrosylation of mitochondria-associated SNO–caspase-3 and promotes apoptotic signaling. **(A)** 10C9 cells were transfected for 3 days with siRNA for TrxR2 or with control RNA before exposure to CH11 monoclonal antibody against Fas (αFas; 50 ng/ml) for 2 hours. The amount of SNO–caspase-3 in a subcellular fraction enriched for mitochondria was evaluated by biotin switch. Results are mean ± SEM of four experiments. *P < 0.05 by analysis of variance (ANOVA). **(B)** 10C9 cells were left untreated or treated with auranofin for 1 hour, followed by Fas receptor stimulation and evaluation of mitochondrial SNO–caspase-3 as in (**A**). Results are the mean ± SEM of three experiments. *P < 0.05 by ANOVA. **(C)** 10C9 cells treated as in (**A**) were lysed in the presence of bVAD-FMK (10 μM), an affinity ligand for active caspase. Cleaved caspase-3 (<20 kD) was assessed by immunoblotting with antibodies against caspase-3 in lysates and after purification of active, bVAD-FMK–bound caspases with streptavidin-agarose (pull-down). Results are representative of three experiments. **(D)** 10C9 cells were transfected with siRNA specific for TrxR1 or TrxR2 or with control RNA. Cells were left untreated or treated with auranofin (1 μM; 1 hour) followed by treatment with anti-Fas (100 ng/ml; 2 hours). Caspase-3 cleavage and activation were assessed as in (**C**). Results (Fas-activated samples) are mean ± SEM of three experiments. *P < 0.05, **P < 0.01 by ANOVA. (**E** and **F**) 10C9 cells were treated with auranofin or transfected with siRNA specific for TrxR1 or TrxR2 as in (**A**) to (**D**). DNA fragmentation was assessed 6 hours after treatment with anti-Fas. Results (Fas-activated samples) are mean ± SEM of four experiments. *P < 0.05 by ANOVA.

caspase-3-mediated apoptosis: caspase-3 activation and DNA fragmentation (*27*). In Fas-stimulated 10C9 cells, both depletion of TrxR2 with siRNA and acute inhibition with auranofin reduced both the amount of cleaved, active caspase-3 [captured with biotin-Val-Ala-Asp(OMe) fluoromethyl ketone (bVAD-FMK)] (fig. 4, C and D, and figs. S6C and S7A) and caspase-3–like activity [cleavage of the tetrapeptide Asp-Glu-Val-Asp (DEVD)] (fig. S6D). In support of these data, activation of caspase-8 (which lies upstream of cytosolic caspase-3) was also diminished by TrxR2 inhibition (fig. S7A). In contrast, depletion of TrxR1 had no appreciable effect on caspase activity (Fig. 4D and fig. S6C). Furthermore, treatment of 10C9 cells with auranofin and knockdown of TrxR2 (but not TrxR1) with siRNA decreased DNA fragmentation by 45% ± 14 (n = 4) (Fig. 4E) and 34% ± 6 (Fig. 4F and fig. S7B), respectively. Although the precise sequence of events subserving transmission of the NO-regulated apoptotic signal from mitochondrial to cytosolic compartments remains to be elucidated fully (further discussed in fig. S7C), our findings suggest that denitrosylation of mitochondria-associated SNO–caspase-3 by Trx2-TrxR2 promotes Fas-induced apoptosis. More generally, inhibition of Trx2-TrxR2 provides a means to selectively manipulate mitochondria-associated caspase.

Concluding remarks. Many SNO-proteins are present constitutively in low amounts, and both *S*-nitrosylation and denitrosylation have been observed after stimulation of multiple classes of receptors (*1*). The present findings provide evidence for a specific enzymatic mechanism of protein denitrosylation, acting in distinct cellular compartments to regulate basal and stimulus-coupled protein denitrosylation, and thus complement the recent description of *S*-nitrosylation and -denitrosylation governed by GSNOR (*11*). Identification of mitochondria- and cytosol-specific denitrosylases not only elucidates the mechanism by which Fas induces denitrosylation of mitochondria-associated caspases selectively but also indicates a function for Trx2 in amplifying apoptotic signals.

Previous studies have raised the idea that Trx1 may promote *S*-nitrosylation of proteins through mechanisms involving nonactive site thiols Cys^{69} or Cys^{73} (*16*, *17*), but these residues are not present in Trx2. Our finding that protein *S*-nitrosylation is augmented by genetic or pharmacologic inhibition of either Trx1-TrxR1 or Trx2-TrxR2 suggests that a major function of the Trx system (involving active-site Cys^{32}) is protein denitrosylation. Most commonly, the Trx system is characterized as playing a highly conserved role in protection from cellular oxidative stress. The present findings indicate that SNO-protein denitrosylation (and thus regulation of NO-based signaling) represents an additional function of the Trx system in metazoan organisms, in which several classes of SNO-regulating enzymes—NO synthases, GSNO reductases, and Trx-TrxRs—are expressed ubiquitously across cell types.

References and Notes

1. D. T. Hess, A. Matsumoto, S. O. Kim, H. E. Marshall, J. S. Stamler, *Nat. Rev. Mol. Cell Biol.* **6**, 150 (2005).
2. J. B. Mannick *et al.*, *Science* **284**, 651 (1999).

3. P. A. Erwin, A. J. Lin, D. E. Golan, T. Michel, *J. Biol. Chem.* **280**, 19888 (2005).
4. J. Hoffmann, J. Haendeler, A. M. Zeiher, S. Dimmeler, *J. Biol. Chem.* **276**, 41383 (2001).
5. J. E. Kim, S. R. Tannenbaum, *J. Biol. Chem.* **279**, 9758 (2004).
6. J. B. Mannick *et al.*, *J. Cell Biol.* **154**, 1111 (2001).
7. Materials and methods are available on *Science* Online.
8. W. Jeong, H. W. Yoon, S. R. Lee, S. G. Rhee, *J. Biol. Chem.* **279**, 3142 (2004).
9. D. R. Arnelle, J. S. Stamler, *Arch. Biochem. Biophys.* **318**, 279 (1995).
10. D. Nikitovic, A. Holmgren, *J. Biol. Chem.* **271**, 19180 (1996).
11. L. Liu *et al.*, *Nature* **410**, 490 (2001).
12. K. Ravi, L. A. Brennan, S. Levic, P. A. Ross, S. M. Black, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 2619 (2004).
13. I. Sliskovic, A. Raturi, B. Mutus, *J. Biol. Chem.* **280**, 8733 (2005).
14. D. A. Stoyanovsky *et al.*, *J. Am. Chem. Soc.* **127**, 15815 (2005).
15. R. Sengupta *et al.*, *Biochemistry* **46**, 8472 (2007).
16. J. Haendeler *et al.*, *Nat. Cell Biol.* **4**, 743 (2002).
17. D. A. Mitchell, S. G. Morton, N. B. Fernhoff, M. A. Marletta, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 11609 (2007).
18. S. R. Jaffrey, H. Erdjument-Bromage, C. D. Ferris, P. Tempst, S. H. Snyder, *Nat. Cell Biol.* **3**, 193 (2001).
19. M. T. Forrester, M. W. Foster, J. S. Stamler, *J. Biol. Chem.* **282**, 13977 (2007).
20. E. S. Arner, M. Bjornstedt, A. Holmgren, *J. Biol. Chem.* **270**, 3479 (1995).
21. S. Gromer, L. D. Arscott, C. H. Williams Jr., R. H. Schirmer, K. Becker, *J. Biol. Chem.* **273**, 20096 (1998).
22. C. H. Lillig, A. Holmgren, *Antioxid. Redox Signal.* **9**, 25 (2007).
23. L. Verdoucq, F. Vignols, J. P. Jacquot, Y. Chartier, Y. Meyer, *J. Biol. Chem.* **274**, 19714 (1999).
24. M. R. Hara *et al.*, *Nat. Cell Biol.* **7**, 665 (2005).
25. S. R. Lee *et al.*, *J. Biol. Chem.* **274**, 4722 (1999).
26. M. Pia Rigobello *et al.*, *J. Inorg. Biochem.* **98**, 1634 (2004).
27. S. Nagata, *Annu. Rev. Genet.* **33**, 29 (1999).
28. The authors thank B. Li and A. Hausladen for technical assistance. M.B. was supported by a Ruth Kirschstein National Research Service Award fellowship from the NIH.

Supporting Online Material

www.sciencemag.org/cgi/content/full/320/5879/1050/DC1
Materials and Methods
Figs. S1 to S7
Tables S1 and S2
References

25 March 2008; accepted 3 April 2008
10.1126/science.1158265

REPORTS

Structural Diversity of Sodium

Eugene Gregoryanz,[1]* Lars F. Lundegaard,[1] Malcolm I. McMahon,[1] Christophe Guillaume,[1] Richard J. Nelmes,[2] Mohamed Mezouar[2]

Sodium exhibits a pronounced minimum of the melting temperature at ~118 gigapascals and 300 kelvin. Using single-crystal high-pressure diffraction techniques, we found that the minimum of the sodium melting curve is associated with a concentration of seven different crystalline phases. Slight changes in pressure and/or temperature induce transitions between numerous structural modifications, several of which are highly complex. The complexity of the phase behavior above 100 gigapascals suggests extraordinary liquid and solid states of sodium at extreme conditions and has implications for other seemingly simple metals.

Some 35 years ago it was argued that as a result of the energy of zero-point oscillations, the ground state of H_2 would be a liquid under sufficient compression (*1*). This idea implies that the melting curve passes through a maximum before decreasing to absolute zero at extreme pressures; such behavior is at odds with our current experience and understanding of melting curves. Although melting curves with negative slopes are known for some metals and molecular compounds, water being the most celebrated example, the negative-slope regions typically exist only at small compressions and over quite limited pressure ranges. The only system in which there are experimental measurements of a decrease in the melting temperature over large pressure ranges to >100 GPa is elemental sodium, which has recently been shown to melt near room temperature around 118 GPa (*2*). Sodium is thus the only known element with an experimentally accessible liquid state at high densities and low temperatures.

Because of their simplicity, light alkali metals have long been used as a paradigm in areas of fundamental physics as diverse as the development of the electron theory of solids (*3*) and the creation of Bose-Einstein condensates (*4*). But in the past decade it has been shown that if the density of lithium or sodium is increased, they start to exhibit phenomena that are counter-intuitive (*5*–*9*). Sodium, in particular, has been found to be a very peculiar system above 100 GPa at 300 K. Relative to the rest of the alkali metals, it has a remarkably wide range of stability of the low-pressure body-centered cubic (bcc) and face-centered cubic (fcc) phases, which extend from ambient pressure up to >100 GPa at 300 K, with a transition between them at 65(1) GPa (here and below, values in parentheses indicate uncertainty in the last significant digit) (*10*). The experimental equation of state gives a bulk modulus above 250 GPa in the 100 GPa pressure range (*10*) (more than half that of diamond at ambient pressure), whereas the shear modulus of the solid near melting must be very small (i.e., close to zero). Calculations of the elastic constants (*11*, *12*), the phonon density of states, and the Debye temperature of Na to >100 GPa (*12*) indicate that the high-frequency region of the spectrum, consisting of longitudinal modes, hardens with pressure while the low-frequency van Hove singularity, corresponding to transversal modes, softens monotonically. As a result, theory predicts that the shear modulus of Na collapses in the fcc phase (*11*, *12*) above 100 GPa. The unusual combination of hard longitudinal phonons and transverse modes softened to zero, at relatively low temperatures (300 K), is likely to give rise to unusual physical behavior of the solid and/or liquid states in the vicinity of the melting minimum at 118 GPa.

Previous synchrotron x-ray powder diffraction studies to >100 GPa did indeed find that Na adopts low-symmetry structures (*2*, *10*, *13*) similar to those seen in Li (*6*). But detailed structural studies of the light alkali metals above 100 GPa (*14*) have become possible only recently with the arrival of new single-crystal experimental techniques (*15*). To investigate the solid phases lying just below the melting minimum, we have exploited sodium's proclivity to form high-quality single crystals under these conditions to perform x-ray single-crystal synchrotron diffraction studies as a function of pressure and temperature to 130 GPa at 500 K, and to 150 GPa at room temperature (*16*).

Within a ±2 GPa range about the minimum at 118 GPa, we have observed six different crystalline phases, five of which have not been observed previously in any other element. Of these five, one is a simple structure (labeled oP8 below) but the other four are very complex, with about 50, 90, 120, and 512 atoms in the unit cell, respectively. Because we obtained high-quality single-crystal data from all the observed phases (Fig. 1), we were able to determine their lattice types and the lattice parameters with confidence, although full solution of these four structures will require further experiments. In two phases we observed twinning, which was easily detectable in the data (*16*). At present, the numbers of atoms per unit cell given below for these structures are estimates to about ±2%, based on the atomic volumes of 9.771(1) Å^3 for the cI16 phase at 115(2) GPa, and 9.445(2) Å^3 for the oP8 phase at 119(2) GPa. To distinguish the numerous phases observed, we label each with its Pearson notation (*17*) to designate the lattice symmetry and number of atoms in the unit cell; for example, the oP8 phase has an orthorhombic primitive lattice and eight atoms in the unit cell.

[1]School of Physics and Centre for Science at Extreme Conditions, University of Edinburgh, Edinburgh EH9 3JZ, UK. [2]European Synchrotron Radiation Facility, BP 220, 38043 Grenoble Cedex, France.

*To whom correspondence should be addressed. E-mail: e.gregoryanz@ed.ac.uk

At 105 GPa, fcc-Na transformed into the cubic cI16 structure (Fig. 2B), as also observed in Li (*6*). Refinements of the new single-crystal data for this structure (Fig. 1A) are in complete agreement with our recently reported structure refinement (*14*). The cI16 phase was stable from 105(1) to 117(2) GPa at 300 K, at which point, in three of the four samples studied, it transformed via a single crystal → single crystal transition into a simple primitive orthorhombic structure with eight atoms per unit cell, oP8 (Fig. 1B and Fig. 2B), and lattice parameters $a = 4.7650(5)$ Å, $b = 3.020(3)$ Å, and $c = 5.251(2)$ Å at 119(2) GPa. The structure has space group *Pnma*. Atoms are on two different 4c sites (x, ¼, z), with atomic coordinates $x = 0.015(1)$ and $z = 0.180(1)$ for one, and $x = 0.164(1)$ and $z = 0.586(1)$ for the other (Fig. 2B), obtained from least-squares structure refinement. The values given above for atomic volumes show that there is a ~2% density increase at the cI16 → oP8 transition.

In the fourth sample, slow compression of the cI16 phase resulted in a single crystal → single crystal transition at 117(2) GPa to an extremely complex primitive monoclinic structure having 512 ± 9 atoms per cell, mP512 (Fig. 1H), with lattice parameters $a = 13.06(3)$ Å, $b = 28.8(1)$ Å, $c = 13.53(3)$ Å, $\beta = 104.9(2)°$, and a volume $V = 4918(20)$ Å^3 at 117(2) GPa (Fig. 2A), just above the transition. To our knowledge, this is the most complex phase ever observed in an element. We note that the intensity distribution in reciprocal space was strongly modulated in a way that is reminiscent of Ga(II) (*18*) and Cs(III) (*19*), which occur close to sharp dips in the Ga and Cs melting curves.

At 300 K, the oP8 phase existed over only a relatively narrow pressure range between 117(2) and 125(2) GPa, above which sodium adopted an incommensurate composite host-guest structure with the same 16-atom tetragonal body-centered host structure observed previously in K(III) and Rb(IV) (*20*, *21*). However, in contrast to K(III), but similar to the disordered phase of Rb(IV) observed between 16.2 and 16.7 GPa (*21*), the one-dimensional guest chains were partially disordered, producing sheets of diffuse scattering that can be seen in Fig. 1D. At 130 GPa, the host and guest lattice parameters are $a_H = a_G = 7.088(3)$ Å, $c_H = 3.483(2)$ Å, and $c_G = 2.11(1)$ Å (Fig. 2B). There are thus $16 \div (2 \times c_H/c_G) = 19.3$ atoms in the unit cell, but for simplicity, we refer to this phase as tI19 (Fig. 1, C and D). In contrast to the disordered phase of Rb(IV), which exists over a pressure range of only 0.5 GPa (*21*), the disordered tI19 phase was stable up to at least 150 GPa at 300 K and up to 500 K at 130 GPa, the highest pressure-temperature conditions reached in this study.

Although we have observed the mP512 phase only once upon compression of the cI16 phase at 300 K (above), we have seen the same phase repeatedly and reproducibly upon recrystallization of liquid Na near the melting minimum at 118 GPa, irrespective of the crystal structure before melting. Two further structures were obtained near 118 GPa via a single crystal → single crystal transition from the mP512 phase as a function of time at 300 K. In one case, mP512 transformed over a period of a few hours into a body-centered tetragonal phase having 50 atoms per unit cell (Fig. 1E) (density considerations give 50 ± 1 atoms per unit cell, but as this structure is body-centered, the number of atoms per cell must be even), with lattice parameters of $a = 7.234(1)$ Å and $c = 9.193(3)$ Å at 117(2) GPa (Fig. 2A). In two other cases, mP512 transformed into a C-centered orthorhombic structure with 120 ± 2 atoms per unit cell, oC120 (Fig. 1F), with lattice parameters $a = 6.65(1)$ Å, $b = 13.17(1)$ Å, and $c = 13.12(2)$ Å at 117(2) GPa (Fig. 2A). On heating near 118 GPa, the cI16, tI50, and oC120 structures were all stable up to the melting curve. The balance of evidence suggests that mP512 may be stable only above 300 K.

At pressures slightly above the melting minimum at 118 GPa, we observed the mP512 phase formed from the melt to transform with time, not to tI50 or oC120, but to another phase, which is triclinic with 90 ± 2 atoms per unit cell, aP90 (Fig. 1G). When melted at a pressure a little above 118 GPa, the liquid crystallized back into this aP90 phase. In this same pressure range, around 119 to 120 GPa, the oP8 phase was found to be stable up to the melting curve, but the liquid again crystallized upon cooling into aP90, with lattice parameters $a = 7.283(5)$ Å, $b = 7.366(5)$ Å, $c = 16.138(9)$ Å, $\alpha = 89.28(3)°$, $\beta = 89.28(3)°$, and $\gamma = 77.96(2)°$ at 119 GPa (Fig. 2A). aP90 is only the third known example of a high-pressure

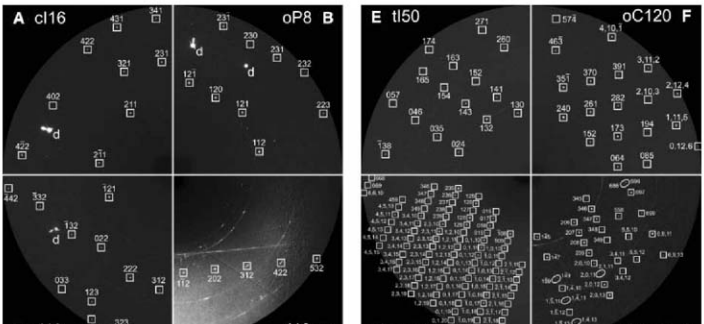


Fig. 1. Eight quadrants of composite diffraction images showing representative data from seven different sodium phases. Squares mark sodium reflections with the given *hkl* indices. Reflections from the diamond anvils are marked by the letter d. (**A** to **C**) Diffraction data from (**A**) cI16-Na at 115(2) GPa collected over a 16° scan range, (**B**) oP8-Na at 119(2) GPa collected over an 11° scan range, and (**C**) tI19-Na at 130 GPa collected over an 11° scan range. (**D**) Stationary long-exposure diffraction image of tI19-Na at ~130 GPa, showing, just above the line of reflections, one line of the diffuse scattering from the guest component of the structure. (**E** to **H**) Diffraction data from (**E**) tI50-Na at 117(2) GPa collected over a 9° scan range, (**F**) oC120-Na at 117(2) GPa collected over a 3° scan range, (**G**) aP90-Na at 119(2) GPa collected over a 4° scan range, and (**H**) mP512-Na at 117(2) GPa collected over a 9° scan range. The five pairs of reflections enclosed in ovals in (**H**) illustrate the close reflection spacing that results from the 29 Å b axis of the mP512 structure. The faint powder rings visible in quadrants (**D**) and (**H**) are from the rhenium gasket. Other small diffraction spots are of a different appearance from those indexed and do not emanate from the sodium sample.

triclinic structure in the elements (*22*). The aP90 structure was observed repeatedly and reproducibly upon recrystallization of liquid Na above 118 GPa near the melting minimum. In all but one case, the aP90 phase obtained from the melt transformed back to oP8 before cooling to 300 K. In the one exception, aP90 was retained to 300 K but, upon gentle warming, it transformed to oP8, which suggests that aP90—like mP512—is stable only above 300 K.

Figure 3 provides a representation of the pressure-temperature regions in which the six phases found close to the melting curve minimum appear to be stable, as suggested by all the observations set out above. The occurrence of six well-crystallized phases (as shown by exceptionally sharp Bragg peaks) in this very small pressure-temperature range, all obtained from single crystal → single crystal transitions, with evidence also of multi-hour kinetics in some cases, suggests a free-energy landscape that has many distinct but shallow minima accessible by only slight changes of pressure and/or temperature.

The existence of such a free-energy landscape at extreme compression raises several questions about the unusual behavior of the melting curve and the liquid and solid states. It was recently suggested that an s-to-p electron transfer mechanism and electronic transitions in the liquid state are responsible for the negative melting slope of Na and the accompanying changes in the solid, leading to the transition of fcc to the cI16 phase and the recovery of a positive melting slope at 100 GPa (*23*). Our results show that this picture is oversimplified and that a full explanation of the unusual melting behavior would need to take into account the actual, extremely complex structures above 100 GPa. Among other things, it is almost certainly the first-order density increase of 2% at the cI16 → oP8 transition that is responsible for the change to a positive melting slope above 118 GPa. Also, the observation of the complex mP512 and aP90 phases next to the melting curve, and the large number of atoms in the unit cells involved, might indicate that the liquid state reflects the structural complexity of the solid below.

One of the hydrogen-related early papers (*24*) proposed the possibility of families of related anisotropic structures in metallic hydrogen, all very close in energy, arising from softening of the transverse phonons. Our results show something very similar being realized in sodium. It appears from calculations that the density-driven changes in the Na phonon density of states soften the shear moduli (*12*) until the solid cannot support the shear stresses, whereupon the liquid state becomes favorable at low temperature and has in its vicinity numerous solid anisotropic structures. It has been shown in other calculations that the Fermi surface of Na remains spherical up to 100 GPa (*25*) and that the whole transverse acoustic phonon branch becomes unstable. As the transverse mode in the Γ-K direction becomes flat and shifts toward lower energies, a simultaneous freezing of many sets of k vectors might take place and give rise to a large number of complex structures (*25*). Interestingly, lithium has strikingly different predicted electronic behavior under pressure. Calculations show that its Fermi surface becomes increasingly anisotropic until it contacts the Brillouin zone boundary and causes strong Fermi surface nesting, which induces a transverse-mode instability that is very well defined in k space (*26*). This instability is thought to be responsible (*26*) for the formation of the low-symmetry complex structures seen in Li at elevated densities (*6*). If the mechanism of the s-to-p transfer at the Fermi surface (*23*) is correct, one might expect a much more pronounced minimum in the melting curve of the lighter Li, with quantum effects playing an important role in the liquid state. However, if the negative melting curve of Na is due to a transverse phonon instability driven by an effect different from s-to-p transfer (*12*), a similar mechanism in lithium would be expected to flatten the Li melting curve but not so far as to make it negative (*25*). It is also interesting to note that Neaton and Ashcroft (*5*) predicted that symmetry breaking in Li under pressure would broaden the core bands and would result in band overlap contributing to pairing states in the valence band. For Na, estimations of the electron density for $Z > 1$ at 118 GPa give $r_s \sim 1.7$ to 1.9 ($r_s a_0 = [3/(4\pi n)]^{1/3}$, where n is the valence electron density and a_0 is

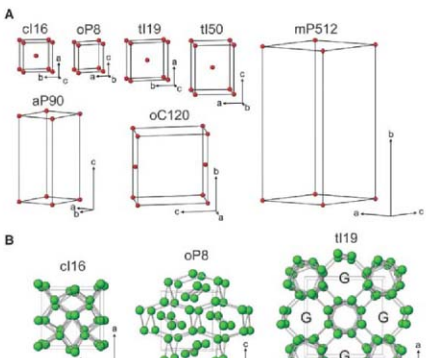


Fig. 2. Structural information for sodium phases stable above 100 GPa. **(A)** Lattices (red points) and unit cells (solid lines) for the seven known phases, with crystallographic axes labeled, all drawn on the same relative scale. **(B)** The atomic arrangement for cI16, oP8, and the host framework of tI19. In tI19, "G" marks channels occupied by disordered chains of guest atoms. The cI16 structure is a distorted 2 × 2 × 2 superstructure of bcc. In the oP8 structure, the atoms lie in planes at y = 1/4 and 3/4 along the b axis. The atoms with $x \sim 0.0$, $z \sim 0.2$ (see text) are shown forming six-sided channels along b, enclosing those with $x \sim 0.2$, $z \sim 0.6$. (Links between atoms are used to help show spatial relationships rather than mark nearest-neighbor contacts.) The actual orientational relationships among the various structures are not known.

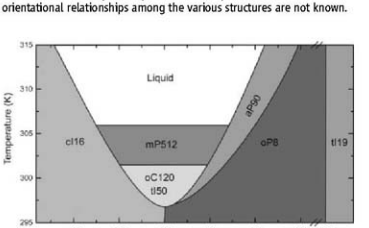


Fig. 3. Observed phases of sodium in the vicinity of the minimum of the melting curve. The melting curve minimum (nominally at 118 GPa) is shown aligned with the cI16-oP8 transition, but this is subject to the ~2 GPa uncertainty on the measured pressures.

the Bohr radius; r_s of Na at 118 GPa is 2.47 for $Z = 1$), and this value approaches the densities at which H_2 is expected to enter the liquid state [see (*27*–*29*) and references therein]. Our preliminary optical measurements show that the reflectivity and the first-order Raman activity of solid Na at 300 K do change above 118 GPa, which suggests a deviation from a metallic state and perhaps indicates a change of valency. However, we have not observed the predicted change of reflectivity versus temperature (*23*) in the liquid state at lower pressures.

Recent experimental measurements on the hydrogen melting curve have suggested a maximum at 130 GPa (*30*) and a possible negative melting curve over a broad pressure range above that, supporting the original idea of a liquid ground state (*1*). Numerous theoretical calculations have studied the properties of hydrogen extensively at extreme compressions, $r_s \leq 1.6$, to explore this possibility further (*27*–*29*). These studies suggested not only superconductivity in the liquid but also novel states of a metallic superfluid (*29*, *31*). The ultimate test of these predictions for H_2 is yet to come, but a similar liquid-solid configuration at high compressions is now established in sodium. Even though the Na atom is heavier and the temperatures involved are higher, one might expect unusual behavior of the Na liquid. Further studies of the liquid state(s) at the minimum of the melting curve are needed to discover whether quantum effects are responsible for the formation of numerous complex solid phases, whether the highly compressed liquid possesses non–simple liquid properties, and what mechanism is responsible for creating such complex behavior in one of the simplest possible systems.

References and Notes

1. E. Brovman, Yu. Kagan, A. Kholas, *Sov. Phys. JETP* **35**, 783 (1972).
2. E. Gregoryanz *et al.*, *Phys. Rev. Lett.* **94**, 185502 (2005).
3. E. Wigner, F. Seitz, *Phys. Rev.* **43**, 0804 (1933).
4. K. Davis *et al.*, *Phys. Rev. Lett.* **75**, 3969 (1995).
5. J. B. Neaton, N. W. Ashcroft, *Nature* **400**, 141 (1999).
6. M. Hanfland *et al.*, *Nature* **408**, 174 (2000).
7. J. B. Neaton, N. W. Ashcroft, *Phys. Rev. Lett.* **86**, 2830 (2001).
8. K. Shimizu *et al.*, *Nature* **419**, 597 (2002).
9. V. V. Struzhkin, M. I. Eremets, W. Gan, H.-k. Mao, R. J. Hemley, *Science* **298**, 1213 (2002); published online 17 October 2002 (10.1126/science.1078535).
10. M. Hanfland, I. Loa, K. Syassen, *Phys. Rev. B* **65**, 184109 (2002).
11. M. Katsnelson *et al.*, *Phys. Rev. B* **61**, 14420 (2000).
12. M. Martinez-Canales, A. Bergara, http://arxiv.org/pdf/0706.2811 (2007).
13. K. Syassen, in *High Pressure Phenomena, Proceedings of the International School of Physics, R. J. Hemley, G. L. Chiarotti, M. Bernasconi, M. L. Ulivi, Eds.* (IOS Press, Amsterdam, 2002), pp. 266–268.
14. M. I. McMahon *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 17297 (2007).
15. L. F. Lundegaard, thesis, University of Edinburgh (2007).
16. See supporting material on *Science* Online.
17. W. B. Pearson, *The Crystal Chemistry and Physics of Metals and Alloys* (Wiley-Interscience, New York, 1972).
18. O. Degtyareva *et al.*, *Phys. Rev. Lett.* **93**, 205502 (2004).
19. M. I. McMahon, R. J. Nelmes, S. Rekhi, *Phys. Rev. Lett.* **87**, 255502 (2001).
20. M. I. McMahon, R. J. Nelmes, U. Schwarz, K. Syassen, *Phys. Rev. B* **74**, 140102(R) (2006).
21. M. I. McMahon, R. J. Nelmes, *Phys. Rev. Lett.* **93**, 055501 (2004).
22. C. Hejny, M. I. McMahon, *Phys. Rev. B* **70**, 184109 (2004).
23. J. Y. Raty, E. Schwegler, S. A. Bonev, *Nature* **449**, 448 (2007).
24. E. Brovman, Yu. Kagan, A. Kholas, *Sov. Phys. JETP* **34**, 1300 (1972).
25. A. Bergara *et al.*, in *Book of Abstracts 21st AIRAPT and 45th EHPRG International Conference*, G. Angilella, P. Pucci, F. Siringo, Eds. (Univ. of Catania, Catania, Italy, 2007), p. 103.
26. A. Rodriguez-Prieto, A. Bergara, V. Silkin, P. Echenique, *Phys. Rev. B* **74**, 172104 (2006).
27. D. Straus, N. Ashcroft, *Phys. Rev. Lett.* **38**, 415 (1977).
28. J. Jaffe, N. Ashcroft, *Phys. Rev. B* **23**, 6176 (1981).
29. E. Babaev, A. Sudbø, N. W. Ashcroft, *Nature* **431**, 666 (2004).
30. E. Gregoryanz *et al.*, *Phys. Rev. Lett.* **90**, 175701 (2003).
31. E. Smargrav, E. Babaev, J. Smiseth, A. Sudbø, *Phys. Rev. Lett.* **95**, 135301 (2005).

32. We thank O. Degtyareva for useful discussions and I. Loa for help with the experiments. Supported by a research grant from the UK Engineering and Physical Sciences Research Council and facilities made available by the European Synchrotron Radiation Facility.

Supporting Online Material

www.sciencemag.org/cgi/content/full/320/5879/1054/DC1
Materials and Methods
Fig. S1
References

28 January 2008; accepted 15 April 2008
10.1126/science.1155715

Inverse Temperature Dependence of Toughness in an Ultrafine Grain-Structure Steel

Yuuji Kimura,* Tadanobu Inoue, Fuxing Yin, Kaneaki Tsuzaki

Materials are typically ductile at higher temperatures and become brittle at lower temperatures. In contrast to the typical ductile-to-brittle transition behavior of body-centered cubic (bcc) steels, we observed an inverse temperature dependence of toughness in an ultrahigh-strength bcc steel with an ultrafine elongated ferrite grain structure that was processed by a thermomechanical treatment without the addition of a large amount of an alloying element. The enhanced toughness is attributed to a delamination that was a result of crack branching on the aligned {100} cleavage planes in the bundles of the ultrafine elongated ferrite grains strengthened by nanometer-sized carbides. In the temperature range from 60° to −60°C, the yield strength was greater, leading to the enhancement of the toughness.

Structural metals, including body-centered cubic (bcc) steels and Ti alloys, usually lose most of their ductility and toughness when the temperature falls below a particular level (*1*). Below the ductile-to-brittle transition temperature (DBTT), a material can fracture in a sharp and sudden way, often leading to serious accidents. The brittle fracture of the Liberty ships in the 1940s is a well-known example of the ductile-to-brittle transition of certain bcc steels (*1*).

The refinement of crystalline grain size is considered to be a key technique for lowering the DBTT for high-strength metals (*2*–*6*). However, strengthening by refining the grain size to 1 μm or less tends to result in low toughness and low tensile ductility at room temperature, although the metals behave in a ductile manner (*3*–*7*).

Delamination, or splitting resulting from anisotropic microstructures such as crystalline grains and second phases that are oriented along the rolling direction (RD), improves the toughness of the metals at low temperatures (*5*, *8*–*11*). Face-centered cubic (fcc) austenitic stainless steel (*10*) and certain Al-Li alloys (*11*), which did not undergo any apparent changes in fracture mode, exhibited marked increases in toughness with decreasing temperature. However, the delamination toughening typically sacrificed toughness in the ductile fracture region for a decrease in DBTT (*4*, *5*, *8*, *9*, *11*).

A martensitic structure consists of thin plate-shaped grains that are highly dislocated, resulting from the diffusionless shear transformation of the fcc austenite phase upon quenching (*2*, *12*). Depending on chemical compositions and heat treatments followed by tempering or aging, the martensitic structure can offer a variety of fine ferrite grain structures interspersed with second-phase particles. Low-alloy steels (≤8 weight % alloying elements) are workhorse materials in most industries because of their low cost and good combinations of strength, ductility, and toughness (*13*). However, low-alloy steels with ultrahigh yield strength (σy) exceeding 1.4 GPa typically exhibit low Charpy V-notch impact energy (vE) of 10 to 40 J at room temperature (*14*–*16*). High-alloy (*17*, *18*) and maraging steels (*19*–*21*) containing large amounts of alloy elements such as Ni and Co are excellent structural materials strengthened by nanometer-sized precipitates; however, they are much more expensive than low-alloy steels. Enhanced toughness can be achieved in the ultrahigh strength level by

National Institute for Materials Science, 1-2-1 Sengen, Tsukuba, Ibaraki 305-0047, Japan.

*To whom correspondence should be addressed. E-mail: kimura.yuuji@nims.go.jp

reducing inclusions and impurities (*18*, *20*). However, even in maraging steels, a vE of 40 J was obtained at a σy of 1.83 GPa (*21*). Such low toughness often restricts their use in certain structural applications.

We describe enhanced toughness at lower temperatures for an ultrahigh-strength low-alloy bcc steel having an ultrafine elongated ferrite grain structure interspersed with nanometer-sized carbides. A hint regarding the microstructural design recently emerged from the deformation of a tempered martensitic structure (*22*). We call this type of thermomechanical treatment "tempforming."

A low-alloy steel containing 0.4% C, 2% Si, 1% Cr, and 1% Mo was studied as a model alloy (*23*). The martensitic structure of the steel was a supersaturated solid solution of C. Tempering the martensitic structure at 500° to 600°C leads to a dispersion of nanometer-sized precipitates in the form of cementite and alloy carbides (*24*) in the martensitic matrix. The steel that was conventionally quenched and tempered at 500°C (the QT sample) exhibited an ultrahigh σy of 1.47 GPa, a tensile strength (σB) of 1.77 GPa, and an adequate elongation (EL) of 10% at room temperature, although its vE was as low as 14 J. However, when the martensitic steel was tempered at 500°C and then deformed to a bar with an equivalent strain of 1.7, namely tempformed at 500°C (the TF sample), the σy, σB, and EL of the steel improved to 1.84 GPa, 1.85 GPa, and 15%, respectively. The vE of the steel improved remarkably to 226 J and was about 16 times higher than that of the QT sample.

The QT sample exhibited a ductile-to-brittle transition in Charpy V-notch impact tests at temperatures from 60° to –100°C, resulting in a reduction of the vE (Fig. 1A). Cracks propagated directly across the central portions of the impact test bars (Fig. 1B), where the fracture mode was observed to be a primary quasi-cleavage (*2*, *16*). Such a ductile-to-brittle transition is well known in ultrahigh-strength bcc steels (*14*–*16*, *19*, *20*). However, the TF samples exhibited delaminations, where cracks branched parallel to the longitudinal direction of the impact test bars, and therefore the crack propagations in the striking direction were substantially blunted (Fig. 1B). Depending on the degree of the delamination, the vE of the TF sample substantially increased from 60°C as the temperature decreased and passed through a maximum in the temperature range of –20° to –60°C, followed by a sharp drop (Fig. 1A). Hence, the TF sample showed an inverse temperature dependence of toughness at lower temperatures, where ultrahigh-strength bcc steels undergo the ductile-to-brittle transition (*14*–*16*, *19*, *20*). At temperatures of –20° to –60°C, some of the specimens did not separate into two pieces. The appearance of these fractures resembles that of bamboo, being much stronger and tougher along the fiber than across it.

Contrary to the QT sample, with a random crystallographic orientation (fig. S1) (*23*), the TF sample had an ultrafine elongated ferrite grain structure with a strong <110> deformation texture that was parallel to the RD (<110>//RD deformation texture) (Fig. 2). The average of transverse linear interceptions for the elongated grains was 260 nm, and spheroidal nanometer-sized carbide particles (≤50 nm) were dispersed in the elongated grain matrix.

The delamination is thought to relax triaxial stress conditions ahead of the advancing crack front and blunt the crack tip when the crack and crack branching planes intersect, thus leading to high impact energy (*25*). The occurrence of delamination requires relatively weak interfaces (*5*, *8*, *10*, *11*, *25*–*27*) or planes (*22*) normal to the notch orientation of the impact test bar. In steels, the crack branching has often been observed to occur along the interfaces between the matrix and inclusions such as elongated MnS (*27*) and carbides (*5*, *8*, *10*) that were in alignment with the RD. For the ultrafine elongated ferrite grain structure with a strong <110>//RD deformation texture, we observed that the fracture mode for the crack branching planes occurred primarily by a quasi-cleavage (fig. S2) (*23*).

The ductile-to-brittle transition occurs when the peak tensile stress in the process zone of a crack tip (σt) exceeds the cleavage fracture stress (σf). The σt scales with the σy and is on the order

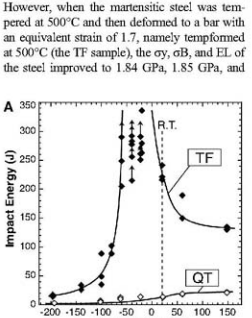


Fig. 1. The results of Charpy V-notch impact tests for the low-alloy steel, conventionally quenched and tempered at 500°C (QT) and tempformed at 500°C (TF). **(A)** Changes in the Charpy V-notch impact energy as a function of testing temperature. R.T., room temperature. **(B)** Appearances of Charpy V-notch impact test specimens after fracture. Data points with upward-pointing arrows indicate that the specimens did not separate into two pieces during the impact test.

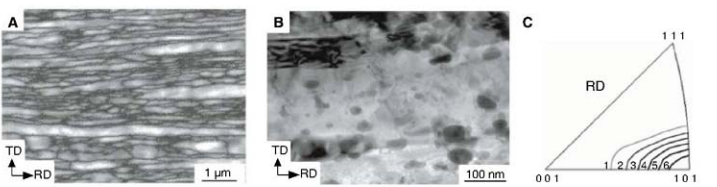


Fig. 2. Ultrafine elongated ferrite grain structure developed in the low-alloy steel that was tempformed at 500°C. **(A)** Image-quality map that was taken by electron backscattering diffraction pattern analysis. Grain boundaries with misorientations above 5° are shown. TD, transverse direction. **(B)** Transmission electron micrograph showing the distribution of nanometer-sized carbides within the elongated grain matrix. **(C)** Inverse pole figure for the RD, showing the evolution of a strong <110>//RD deformation texture.

of three to five times higher than the σy (*2*). The thermal increment in the σy with decreasing temperature eventually causes the cleavage fracture of bcc steel. The QT sample lost most of the tensile ductility when the σy increased to 1.78 GPa at –196°C (Fig. 3). The σy of the TF sample also markedly increased as the temperature decreased. Hence, this indicates that the delamination may be more pronounced in the response to the cleavage fracture at lower temperatures, leading to the onset of the inverse temperature dependence of toughness (Fig. 1). A similar inverse temperature dependence of toughness has already been observed in the vicinity of 200°C for a 0.2% C, 3% Ni, 3% Mo bcc steel processed by a thermomechanical treatment that involved the deformation of metastable austenite, followed by quenching and tempering (*26*). However, the fracture proceeded easily without much plastic deformation below 200°C, although the tendency for crack branching along the boundaries of elongated prior-austenite grains in the tempered martensitic structure was strong. As a result, the vE was reduced to 33 J at room temperature.

Bcc iron cleaves on {100} planes, and the coherence length on {100} corresponds to the cleavage crack length in steel (*2*). In the elongated grain structure with a strong <110>//RD deformation texture, lots of {100} cleavage planes exist along the RD, whereas the planes that are normal to the RD consist of {110} ductile orientations. The coherence length on {100} is long because of the elongated grain. This structure produces the condition where the crack can run along the longitudinal direction rather than the transverse direction.

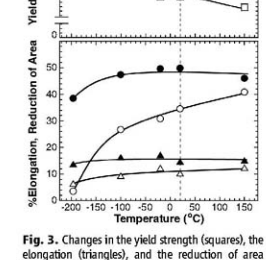


Fig. 3. Changes in the yield strength (squares), the elongation (triangles), and the reduction of area (circles) of the low-alloy steel that was subjected to conventional quenching and tempering at 500°C (QT, open symbols) and tempforming at 500°C (TF, solid symbols) as a function of testing temperature.

Decreasing the transverse grain size to submicrometer sizes leads to a substantial strengthening in the elongated grain structure (*28*). According to the Yoffee diagram, refining the grain size increases both the σt and the σf; however, the effect on the σf is ordinarily larger with a decrease in DBTT (*2*). Hence, the refinement of the transverse grain size can provide more ductile and tougher planes normal to the longitudinal direction of the elongated grain structure at lower temperatures. This is supported by the fact that a good tensile ductility was retained by the TF sample even at –196°C, where the σy increased to 2.29 GPa (Fig. 3). The EL and the reduction of area were 14 and 39%, respectively.

The dispersion of nanometer-sized precipitates may contribute to the strengthening of the steel without lessening ductility and toughness (*17*, *18*). Because the ductile fracture occurred by the growth and coalescence of voids typically nucleated at second-phase particles, the volume fraction (f) of second particles should be minimized. Meanwhile, the strain-hardening rate for the ultrafine-grained steel has been observed to be proportional to the dispersion parameter $(f/d)^{0.5}$, where d is the average diameter of the particles (*29*). Hence, the distribution of a small volume fraction of nanometer-sized particles is effective for improving the tensile ductility, especially uniform elongation in the ultrafine ferrite grain structure (*30*).

To verify our hypothesis, the TF sample was annealed at 700°C for 60 min (the TFA sample). The coarsening of the microstructure (fig. S3) (*23*) resulted in the reduction of the vE as well as the σy and the properties were similar to those of the band of low-alloy tempered martensitic steels (Fig. 4). Therefore, it may be concluded that the combination of the ultra grain refinement, the dispersion of nanometer-sized particles, and the texture control can lead to ultrahigh strength with enhanced toughness as well as ductility in bcc steels at lower temperatures. Our microstructural approach without adding a large amount of alloy elements surpasses the toughness of maraging steel, which is the strongest structural material presently used (*19*–*21*). The strategy described above is expected to be widely applicable in the production of ultrahigh-strength material bar and parts such as bolts and shafts, where ductile-to-brittle transitions occur.

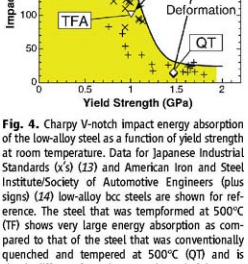


Fig. 4. Charpy V-notch impact energy absorption of the low-alloy steel as a function of yield strength at room temperature. Data for Japanese Industrial Standards (x's) (*13*) and American Iron and Steel Institute/Society of Automotive Engineers (plus signs) (*14*) low-alloy bcc steels are shown for reference. The steel that was tempformed at 500°C (TF) shows very large energy absorption as compared to that of the steel that was conventionally quenched and tempered at 500°C (QT) and is clearly different from the general trend of the commercial low-alloy steels. Degradation of the impact energy as well as the yield strength occurs as a result of the annealing at 700°C for the TF steel (TFA).

References and Notes

1. W. T. Becker, in *ASM Handbook* (American Society for Metals (ASM) International, Novelty, OH, 2002), vol. 11, pp. 557–745.
2. J. W. Morris Jr., C. S. Lee, Z. Guo, *ISIJ Int.* **43**, 410 (2003).
3. S. Takaki, K. Kawasaki, Y. Kimura, *J. Mater. Process. Technol.* **117**, 359 (2001).
4. N. Tsuji, S. Okuno, Y. Koizumi, Y. Minamino, *Mater. Trans.* **45**, 2272 (2004).
5. R. Song, D. Ponge, D. Raabe, *Acta Mater.* **53**, 4881 (2005).
6. V. V. Stolyarov, R. Z. Valiev, Y. T. Zhu, *Appl. Phys. Lett.* **88**, 041905 (2006).
7. A. A. Karimpoor, K. T. Aust, U. Erb, *Scr. Mater.* **56**, 201 (2007).
8. P. Shanmugam, S. D. Pathak, *Eng. Fract. Mech.* **53**, 991 (1996).
9. B. Mintz, W. B. Morrison, *Mat. Sci. Tech.* **23**, 1346 (2007).
10. W. Zhou, N. L. Loh, *Scr. Mater.* **34**, 633 (1996).
11. K. T. Venkateswara Rao, W. Yu, R. O. Ritchie, *Metall. Trans. A* **20A**, 485 (1989).
12. S. Morito, H. Tanaka, R. Konishi, T. Furuhara, T. Maki, *Acta Mater.* **51**, 1789 (2003).
13. S. Nishijima, A. Ishii, K. Kanazawa, S. Matsuoka, C. Masuda, *NRIM Fatigue Data Sheet Technical Document No. 5* (National Research Institute for Metals, Tokyo, 1989).
14. T. V. Philip, T. J. McCaffrey, in *Metals Handbook* (ASM International, Novelty, OH, ed. 10, 1990), vol. 1, pp. 430–448.
15. Y. Tomita, *Int. Mater. Rev.* **45**, 27 (2000).
16. Y. Tomita, *Metall. Trans. A* **22A**, 1093 (1991).
17. J. L. Maloney, W. M. Garrison Jr., *Acta Mater.* **53**, 533 (2005).
18. L. E. Iorio, W. M. Garrison Jr., *Metall. Mater. Trans. A* **37**, 1165 (2006).
19. G. W. Tuffnell, R. L. Cairns, *Trans. ASM* **61**, 798 (1968).
20. G. J. Spaeder, *Metall. Trans.* **1**, 2011 (1970).
21. K. Rohrbach, M. Schmidt, in *Metals Handbook* (ASM International, Novelty, OH, ed. 10, 1990), vol. 1, pp. 793–800.
22. Y. Kimura, T. Inoue, F. Yin, O. Sitdikov, K. Tsuzaki, *Scr. Mater.* **57**, 465 (2007).
23. See the supporting material on *Science* Online.
24. G. Krauss, *Steels: Heat Treatment and Processing Principles* (ASM International, Novelty, OH, 1990), pp. 205–261.
25. D. W. Kum, T. Oyama, J. Wadsworth, O. D. Sherby, *J. Mech. Phys.* **31**, 173 (1983).
26. A. J. McEvily Jr., R. H. Bush, *Trans. ASM* **55**, 654 (1962).
27. C. M. Yen, C. A. Stickels, *Metall. Trans.* **1**, 3037 (1970).
28. A. Belyakov, Y. Kimura, Y. Adachi, K. Tsuzaki, *Mater. Trans.* **45**, 2812 (2004).
29. A. Ohmori, S. Torizuka, K. Nagai, *ISIJ Int.* **44**, 1063 (2004).
30. Y. Sakai, M. Ohtaguchi, Y. Kimura, K. Tsuzaki, *Proceedings Ultrafine Grained Materials*, R. S. Mishra *et al.*, Eds. (The Minerals, Metals and Materials Society, Warrendale, PA, 2000), pp. 361–370.

31. We thank S. Kuroda, M. Fujiwara, Y. Ueno, T. Hibaru, and S. Iwasaki for the materials processing; K. Nakazato for the Charpy impact test; and Y. Hirota, A. Sakurai, E. Motoki, and I. Sakamaki for their experimental assistance in the microstructural observations. This work is a fruit of the Ultra-Steel Project, which began in 1997 and was completed in March 2006 at the National Institute for Materials Science in Japan.

Supporting Online Material
www.sciencemag.org/cgi/content/full/320/5879/1057/DC1
Materials and Methods
SOM Text
Figs. S1 to S3
References

5 February 2008; accepted 14 April 2008
10.1126/science.1156084

Dislocation-Driven Nanowire Growth and Eshelby Twist

Matthew J. Bierman,[1*] Y. K. Albert Lau,[1*] Alexander V. Kvit,[2] Andrew L. Schmitt,[1] Song Jin[1]†

Hierarchical nanostructures of lead sulfide nanowires resembling pine trees were synthesized by chemical vapor deposition. Structural characterization revealed a screwlike dislocation in the nanowire trunks with helically rotating epitaxial branch nanowires. It is suggested that the screw component of an axial dislocation provides the self-perpetuating steps to enable one-dimensional crystal growth, in contrast to mechanisms that require metal catalysts. The rotating trunks and branches are the consequence of the Eshelby twist of screw dislocations with a dislocation Burgers vector along the ⟨110⟩ directions having an estimated magnitude of 6 ± 2 angstroms for the screw component. The results confirm the Eshelby theory of dislocations, and the proposed nanowire growth mechanism could be general to many materials.

In the burgeoning field of nanoscience, a major ambition is to synthesize nanoscale building blocks of arbitrary dimensions, morphologies, and materials of increasing complexity. One-dimensional (1D) nanowire materials, in particular, have already found many applications in nanoelectronics, nanophotonics, and biotechnology (*1*, *2*). To break the symmetry of bulk crystals and enable the anisotropic 1D crystal growth of inorganic nanowires, the well-known vapor-liquid-solid (VLS) growth method uses metal nanoparticles that form low–melting point eutectic alloys with the targeted materials to serve as the catalytic seeds for 1D anisotropic growth (*3*, *4*). Except for direct vapor-solid growth (*5*), most nanowire-formation mechanisms, including solution-liquid-solid growth (SLS) (*6*) and variants of VLS such as vapor-solid-solid growth (*7*), require the use of catalytic nanoparticles, either added intentionally or generated in situ, to enable the 1D anisotropic crystal growth. "Treelike" or hyperbranched nanostructures have also been reported, but they all rely on multiple applications of metal catalysts with subsequent VLS (*8*, *9*) or SLS (*10*) growth steps.

We suggest a nanowire growth mechanism that does not depend on catalysts but instead is driven by an axial screwlike dislocation along the length of the nanowire. It results in hierarchical lead sulfide (PbS) nanostructures of pine tree morphology when combined with a slower in situ VLS branching nanowire growth. The geometrical features of the resulting structures can be quantitatively understood with the simple elasticity theory of dislocations.

The nanostructures of PbS are synthesized via chemical vapor deposition with $PbCl_2$ and elemental sulfur as precursors under argon flow with a co-flow of H_2 at atmospheric pressure and with temperatures between 600° and 650°C (*11*). Typical synthesis conditions involve reactions at 600°C for 15 min under 150 standard cubic centimeters per minute (sccm) of argon flow and 900 torr pressure with the hydrogen flow at 1.8 sccm for the first 1 min and 1.0 sccm for the remaining 14 min. Even though the synthetic procedure is similar to that for the hyperbranched PbS nanowires (see examples in figs. S2 and S3) (*12*), the nanowire growth appears to be driven by different mechanisms. The key difference between the growth of pine trees and the growth of hyperbranched nanowires (*12*) and other previously reported PbQ (Q is S, Se, or Te) nanowire growth (*13*, *14*) is the hydrogen flow profile. The optimized reactions (*11*) reproducibly yield many intricate treelike PbS nanowire structures over large areas (1 to 2 cm^2) on the growth substrate, as revealed by scanning electron microscopy (SEM) (Fig. 1 and fig. S1; also see fig. S4 for phase identification). These trees have trunks that are up to hundreds of micrometers in length and branches that are commonly tens of micrometers long. Individual wires grow consistently along the ⟨100⟩ crystallographic directions and their diameters range from 40 to 350 nm. Closer exam-

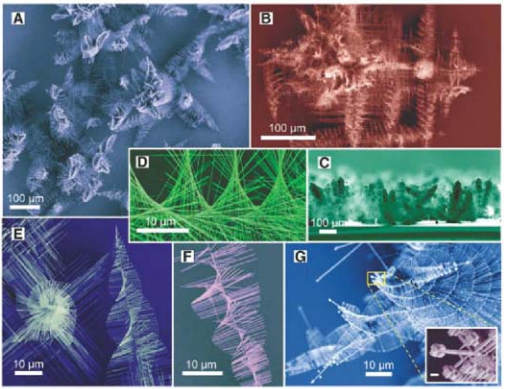


Fig. 1. SEM micrographs of PbS pine tree nanowires. **(A)** Overview of dense forest of many nanowire trees. **(B)** Tree clusters showing epitaxial growth along ⟨100⟩ directions. **(C)** Side view of growth substrate showing forest growth. **(D** to **F)** High-magnification views of trees highlighting the twisting (Eshelby twist) of the central trunk and helical rotating branches, with **(E)** further illustrating branch epitaxy on the tree trunk and **(F)** showing a tree with fewer branches. **(G)** An example of "tree-on-tree" morphology that can be occasionally observed. (Inset) A magnified view of the tips of nanowires after synthesis highlighting the cubes that sometimes decorate the tips. The inset scale bar is 200 nm. The images are false colored.

[1]Department of Chemistry, University of Wisconsin–Madison, 1101 University Avenue, Madison, WI 53706, USA. [2]Materials Science Center, University of Wisconsin–Madison, 1509 University Avenue, Madison, WI 53706, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: jin@chem.wisc.edu

ination of these nanostructures, particularly those with less dense branching (Fig. 1, D to F), reveals that each tree has four sets of epitaxial branches that are perpendicular to the trunk and the neighboring branches and rotate around the trunk in a helical staircase fashion. The pitch of the rotations ranges from 16 to 220 μm and can vary down the length of a single tree. Right- or left-handed rotating trees have roughly equal probability of occurrence (measured ratio: 107:126) (*11*). The rotating branches become progressively shorter from the base to the tip of the trees, resulting in conelike envelopes that enforce the tree morphology. The combination of the helical rotation and the regular length progression of the evenly spaced branches leads to beautiful curves formed by the branch tips, which are further accented when the tips are decorated with PbS cubes in some syntheses (see an example in the inset of Fig. 1G). These trees can grow both upward from the growth substrate and horizontally to the substrate from some nucleation clusters, creating a dense copse of freestanding nanowire trees, which often resembles a forest when viewed from the side of the growth substrates (Fig. 1C). Within a common cluster, the trees occasionally grow epitaxially as evidenced by the perpendicular or parallel orientations between the trees (Fig. 1B). Tree structures sometimes coexist with or grow off from hyperbranched nanowire clusters (Fig. 1E and fig. S2). Occasionally, multiple levels of "tree-on-tree" morphology can be observed (Fig. 1G).

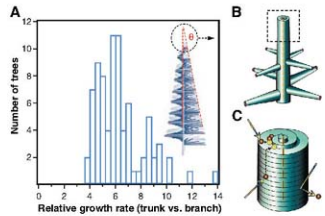


Fig. 2. Distinctive difference in the growth rates of trunk and branch nanowires and the proposed dislocation-driven nanowire growth in the trunk of tree structures. **(A)** Approximate relative ratios of growth rates between trunk and branch nanowires that are calculated as cotangents of the cone angles (θ as illustrated in the inset) of the outer envelopes of 80 individual trees. **(B)** Dramatized scheme of the magnified tip of a tree structure highlighting the combined faster dislocation-driven trunk nanowire growth and slower VLS-driven branched nanowire growth. **(C)** A simplified scheme illustrating that the self-perpetuating steps of a screw dislocation spiral at the tip of a trunk can enable 1D crystal growth of nanowires.

Further scrutiny of the distinct morphology and large length difference between the trunks and branches of these pine tree nanowires suggests that the trunks grow faster than the branches. It was previously suggested that hyperbranched PbS nanowires are grown via a self-catalyzed VLS mechanism: Lead itself serves as the eutectic catalyst and is consumed or evaporated during growth (*15*), thereby limiting the length of wire growth (*12*, *13*). Although no lead catalyst caps were observed at the tips of either the trunk or branch nanowires after growth, a similar length limit was sometimes also observed for the tree branches (fig. S5), suggesting that the branch growth is likely VLS driven. Successive generations of hyperbranched nanowires (*12*, *13*) grow at a similar rate and lead to more isotropic "cubic" morphology (figs. S2 and S3). In contrast, the pine trees have steep cone angles that do not change over the duration of growth for a given tree, suggesting that the trunk nanowires grow much faster than the branches in the tree structures. The cone angles of the outer envelopes of the trees are dictated by the relative ratios between the fastest growth rates of the trunks and branches and do not depend on the actual length and possible delays in nucleation events. To quantitatively represent the growth rate difference between the trunks and branches, the cotangent of the cone angle (θ) of 80 tree envelopes is measured and ranges from 4 to about 10 (Fig. 2A). This distinct morphology and growth rate difference contrasts with the hyperbranched nanowires under similar reaction conditions, suggesting a different growth process for the trunks of pine tree nanowires.

We propose that the growth of the trunk nanowires is driven by the screw dislocation component of an axial dislocation along the length of the nanowire, providing a continuous growth front for 1D crystal growth. The self-perpetuating steps of a screw dislocation provide facile spiral growth fronts when the supersaturation is lower than what is required for crystal growth on perfect crystal facets, and this is known as Frank's mechanism for crystal growth (*16*, *17*). When the supersaturation is low, only fast crystal growth at the self-perpetuating steps of a screw dislocation

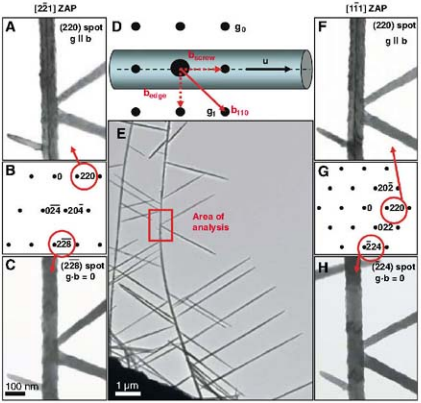


Fig. 3. Diffraction contrast TEM imaging of the dislocation in the tree trunk. **(A** to **C)** TEM images along the [2$\bar{2}$1] zone axis under the strong two-beam conditions. **(A)** represents strong diffraction contrast, and **(C)** represents invisibility conditions as highlighted in the zone axis pattern (ZAP) **(B)**. **(D)** Schematic superposition of real and reciprocal space of a dislocation-containing nanowire along the [001] zone axis illustrating the Burgers vector relationship. Because the line dislocation vector **u** is [100], directions of screw and edge character Burgers vectors, $\mathbf{b}_{screw}$ and $\mathbf{b}_{edge}$, are known. As the Burgers vector is determined to be $\mathbf{b}_{110}$, the high-contrast results at $\mathbf{g}_1$ are due to the $\mathbf{g}_1$**||b** relationship and the perpendicular $\mathbf{g}_0$-beam results in the invisibility criterion. **(E)** Low-magnification TEM image showing the tree and area analyzed. **(F)** to **(H)** display [1$\bar{1}$1] zone axis TEM under the strong two-beam conditions. (F) represents **g||b** conditions, (H) represents invisibility conditions as highlighted in the zone axis diagram (G).

spiral is possible, whereas growth on the crystalline side walls is suppressed (Fig. 2C). This breaks down the symmetry and drives the 1D anisotropic crystal growth without catalysts. This dislocation-driven growth was proposed in the 1950s by Sears to explain the formation of micrometer-diameter metal "whiskers" (*18*, *19*), which predates the VLS whisker growth. However, starting from the original Wagner and Ellis VLS work (*4*, *20*), much effort has been undertaken to rule out crystal dislocations as the driving force for the 1D anisotropic growth. Since then, little has been mentioned about the role of dislocation defects in whiskers (*1*, *21*) (and now nanowires).

We confirm the presence of screw dislocations in the trunks of these tree structures using diffraction contrast transmission electron microscopy (TEM) under the strong two-beam conditions (*22*). Diffraction contrast TEM is a powerful technique to image dislocations in crystals that relies on additional electron diffraction due to the bending of atomic planes near the dislocation core. If an image is reconstructed from specific reciprocal space diffraction spots (**g**) that are selected by a physical aperture, these additional diffracted electrons create a visible contrast around the dislocation. However, certain diffraction spots (**g**) with specific orientations to the Burgers vector of the dislocation (**b**) produce no dislocation contrast—the "invisibility criterion" (*11*). TEM sample preparation proved to be difficult due to the need to preserve the tree morphology during transfer, while also avoiding trees with too many branches that would obstruct the view and prevent the observation of a dislocation. After experimenting with many different transfer methods, we found micromanipulation to be the only technique that allows the effective transfer of individual trees onto TEM grids while preserving their complex and fragile morphology (fig. S6 and movie S1) (*11*). Great care was taken to avoid excessive mechanical force, which can result in the dislocation being worked out along the slip planes. In trees that clearly preserve the twisting structures under microscope observation (such as in Fig. 3E), dark lines running the entire length of tree trunks representing high dislocation contrast were observed under TEM. This is shown for $\mathbf{g} = (220)$ in Fig. 3A along the $[2\bar{2}1]$ zone axis, in Fig. 3F along the $[1\bar{1}1]$ zone axis, and in fig. S7 with a more complete mapping. However, no dislocations were observed in the branches of any tree investigated. No dislocations were observed in hyperbranched nanowires (fig. S9), with more than 20 samples having been examined. These observations are consistent with the suggestion that the nanowire trunks in the trees are driven by dislocation, whereas the branches of the trees (and the hyperbranched nanowires) grow via a slower VLS process (Fig. 2B).

We have determined the dislocation Burgers vector (**b**) to be along the [110] direction. This detailed diffraction contrast TEM analysis requires finding two noncollinear diffraction spots (**g** beams) in reciprocal space that satisfy the invisibility criterion. The dislocation Burgers vector is along the direction of the cross product of these two **g** vectors. The tree structure shown in Fig. 3E has been analyzed under the strong two-beam conditions, as illustrated schematically in Fig. 3D. The same segment of this tree was tilted to the $[2\bar{2}1]$ zone axis (Fig. 3, A to C) and the $[1\bar{1}1]$ zone axis (Fig. 3, F to H), respectively. The image with the (220) diffraction spot (Fig. 3A) shows high dislocation contrast (corresponding to the $\mathbf{g}\|\mathbf{b}$ contrast maximum), while the dislocation meets the invisibility criterion under the perpendicular $(2\bar{2}8)$ spot (Fig. 1C). Similarly, along the $[1\bar{1}1]$ zone axis, the (220) diffraction spot shows high contrast (Fig. 3F) while the perpendicular $(2\bar{2}4)$ spot (Fig. 3H) meets the invisibility criterion. Therefore, taking the cross product of the $(2\bar{2}8)$ and $(2\bar{2}4)$ vectors shows that the Burgers vector is along the [110] direction. Electron diffraction patterns of the area analyzed are shown in fig. S8. It is known that the Burgers vector of the most stable dislocations in rock salt (face-centered cubic) crystals is along ⟨110⟩, and this has been previously observed in bulk PbS crystals (*23*, *24*). Because the dislocation line direction (**u**) is along the [100] nanowire growth direction, the [110] Burgers vector represents a mixed dislocation: a screw dislocation component along the [100] (or $[\bar{1}00]$) direction mostly responsible for driving the nanowire growth, and an edge dislocation component along the [010] (or $[0\bar{1}0]$) direction, whose role in promoting crystal growth (*25*) is not clear but cannot be ruled out completely at present.

What, then, is the reason for the helical rotation of branches on the screw dislocation–driven nanowire trunks? All dislocations create strain (and hence stress) within the otherwise perfect crystalline lattice. Using elasticity theory, Eshelby has shown that in a finite cylindrical rod containing an axial screw dislocation at the center, the stress field created by the dislocation exerts a torque at the free ends of the rod, resulting in a twist of the rod along the axial direction (Fig. 4A) (*23*, *26*). This "Eshelby twist" is mathematically expressed as:

$$\alpha = \frac{b}{\pi R^2} \qquad (1)$$

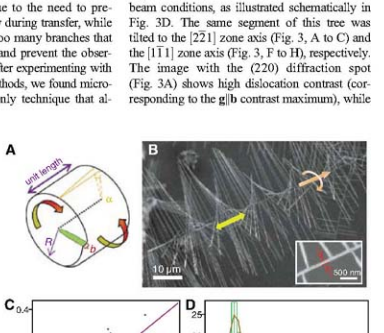


Fig. 4. Analysis of the Eshelby twists in tree nanostructures. **(A)** Schematic representation of the forces and resulting crystal displacement due to a screw dislocation. **(B)** SEM images of a tree illustrating the measurement of twist (a quarter of the pitch measured) and the measurement of diameter (inset), which was converted to radius for calculation. **(C)** Scatterplot of twists measured from 247 spots on 90 individual trees against their inversed cross-sectional areas $[(\pi R^2)^{-1}]$. The red line is a least-squares fit through the data whose slope (6 Å) is the magnitude of the screw component of the Burgers vector. **(D)** Histogram of the calculated Burgers vectors for each data point shown in (C) with a Gaussian fit to the data. The Gaussian peak is centered at 6 Å with a standard deviation of 2 Å.

where α is the twist of the lattice in radians per unit length, R is the radius of the cylinder, and b is the magnitude of the screw component of the Burgers vector (*27*). Attempts to observe the Eshelby twist in micrometer-scale whiskers were made in the late 1950s, but the results were often inconclusive (*28*, *29*). The Eshelby twist is readily observed in the tree nanowires because the $1/R^2$ dependence makes the twist much more pronounced at the nanoscale compared to the micrometer-sized whiskers, and because the overgrowth of epitaxial branching nanowires allows easy visualization of the twist. This allows a direct measurement of Eshelby twists and a simple estimate of the magnitude of the Burgers vector screw component.

As illustrated in Fig. 4B, SEM images can be examined to determine both the radius of a trunk nanowire and also its twist by tracking the periodic repeat of the branches (a quarter of the pitch is actually measured because of the four orthogonal epitaxial branches). The Eshelby twists (α) as a function of the inverse cross-sectional areas $[(\pi R^2)^{-1}]$ of the nanowires were

measured at 247 points on 90 individual trees from 16 synthetic batches (Fig. 4C). To extract the magnitude of the Burgers vector, a line can be fit to the data as plotted, with the slope representing b from Eq. 1 above. This can be more directly seen in a histogram of the calculated Burgers vectors (Fig. 4D). A Gaussian fit to these data yields the average magnitude of the screw component of the Burgers vector $b = 6 \pm 2$ Å. Because the Burgers vector direction is confirmed to be [110] by TEM, a 6 Å screw component of the Burgers vector (the projection of **b** onto the dislocation line **u** [100]) is approximately equal to the lattice constant of PbS, $a = 5.94$ Å. It is known that smallest **b** allowable is the shortest lattice translation vector in a material (*23*), which in the case of rock salt crystals is ½⟨110⟩, whose screw component is ½⟨100⟩ (half the lattice constant a). Given various sources of errors in this estimate (*27*), it is satisfying to see that no data were observed substantially below the theoretical minimal vector, and the average estimated b value of twice the minimal theoretical length is reasonable. Additionally, theory predicts that left-handed dislocation spirals lead to right-handed Eshelby twists and vice versa (*23*, *26*); therefore, the equal probability of twist handedness implies equal probability of Burgers vector sense (sign).

The observation of Eshelby twist in these pine tree nanowires is a clear demonstration and validation of Eshelby's theory on dislocations. The results also provide evidence for a catalyst-free nanowire growth mechanism driven by axial screw dislocations and imply that VLS and screw dislocation-driven nanowire growth can coexist. Because of the distinct morphology difference from the hyperbranched nanowires, it is unlikely that the dislocation is a result of cool-down or other postgrowth perturbation. Although some general discussions on the origins of dislocations exist (*23*, *30*), an experimentally observed mechanistic understanding is currently lacking. We suggest that this dislocation-driven nanowire growth mechanism proposed for PbS trees is likely general to and is underappreciated in the synthesis of 1D nanostructures, particularly in cases where the growth mechanism is inconclusively explained and especially when free of catalysts. Besides the analogous PbSe for which we have found preliminary evidence of similar growth phenomena, the dislocation-driven nanowire growth mechanism is likely to occur in materials that are prone to have screw dislocations, such as SiC, GaN, ZnO, and CdS, both in vapor-phase growth and in solution-phase synthesis. However, we caution that postgrowth mechanical perturbation could work the dislocation out of the nanowires, and one might not be able to observe dislocations in the final nanowire products if samples are not handled properly.

References and Notes

1. Y. Xia *et al.*, *Adv. Mater.* **15**, 353 (2003).
2. C. M. Lieber, Z. L. Wang, *MRS Bull.* **32**, 99 (2006).
3. A. M. Morales, C. M. Lieber, *Science* **279**, 208 (1998).
4. R. S. Wagner, W. C. Ellis, *Appl. Phys. Lett.* **4**, 89 (1964).
5. Z. W. Pan, Z. R. Dai, Z. L. Wang, *Science* **291**, 1947 (2001).
6. T. J. Trentler *et al.*, *Science* **270**, 1791 (1995).
7. A. I. Persson *et al.*, *Nat. Mater.* **3**, 677 (2004).
8. K. A. Dick *et al.*, *Nat. Mater.* **3**, 380 (2004).
9. D. Wang, F. Qian, C. Yang, Z. H. Zhong, C. M. Lieber, *Nano Lett.* **4**, 871 (2004).
10. A. Dong, R. Tang, W. E. Buhro, *J. Am. Chem. Soc.* **129**, 12254 (2007).
11. Materials and methods are available as supporting material on *Science* Online.
12. M. J. Bierman, Y. K. A. Lau, S. Jin, *Nano Lett.* **7**, 2907 (2007).
13. J. Zhu *et al.*, *Nano Lett.* **7**, 1095 (2007).
14. M. Fardy, A. L. Hochbaum, J. Goldberger, M. M. Zhang, P. Yang, *Adv. Mater.* **19**, 3047 (2007).
15. J. B. Hannon, S. Kodambaka, F. M. Ross, R. M. Tromp, *Nature* **440**, 69 (2006).
16. W. K. Burton, N. Cabrera, F. C. Frank, *Nature* **163**, 398 (1949).
17. W. K. Burton, N. Cabrera, F. C. Frank, *Philos. Trans. R. Soc. London A* **243**, 299 (1951).
18. G. W. Sears, *Acta Metall.* **1**, 457 (1953).
19. G. W. Sears, *Acta Metall.* **3**, 361 (1955).
20. R. S. Wagner, W. C. Ellis, K. A. Jackson, S. M. Arnold, *J. Appl. Phys.* **35**, 2993 (1964).
21. D. R. Veblen, J. E. Post, *Am. Mineral.* **68**, 790 (1983).
22. D. B. Williams, C. B. Carter, *Transmission Electron Microscopy: A Textbook for Materials Science* (Plenum, New York, 1996), chaps. 22 and 25.
23. J. P. Hirth, J. Lothe, *Theory of Dislocations* (McGraw-Hill, New York, 1968).
24. A. Foitzik, W. Skrotzki, P. Haasen, *Physica Status Solidi A* **121**, 81 (1990).
25. E. Bauser, H. Strunk, *J. Cryst. Growth* **51**, 362 (1981).
26. J. D. Eshelby, *J. Appl. Phys.* **24**, 176 (1953).
27. Edge dislocations do not produce such distortions. Mixed dislocations can be evaluated as separate screw and edge components. When the cross section of the nanowire is not circular (which is often the case), the cross-section area can be used together with a small correction factor for this calculation. When the dislocation line is not at the center of the cylinder, a small correction factor is applied.
28. R. D. Dragsdorf, W. W. Webb, *J. Appl. Phys.* **29**, 817 (1958).
29. G. W. Sears, *J. Chem. Phys.* **31**, 53 (1959).
30. F. R. N. Nabarro, *Theory of Crystal Dislocations* (Oxford Univ. Press, London, 1967).

31. S. J. thanks NSF (CAREER DMR-0548232), Research Corporation Cottrell Scholar Award, DuPont Young Professor Grant, and 3M Nontenured Faculty Award for support. M.J.B. was partially supported by an Air Products Fellowship. We thank R. Selinsky for assistance with the illustrations in Fig. 2.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1157131/DC1
Materials and Methods
Figs. S1 to S9
References
Movie S1

29 February 2008; accepted 16 April 2008
Published online 1 May 2008;
10.1126/science.1157131
Include this information when citing this paper.

Detection of Silica-Rich Deposits on Mars

S. W. Squyres,[1]* R. E. Arvidson,[2] S. Ruff,[3] R. Gellert,[4] R. V. Morris,[5] D. W. Ming,[5] L. Crumpler,[6] J. D. Farmer,[3] D. J. Des Marais,[7] A. Yen,[8] S. M. McLennan,[9] W. Calvin,[10] J. F. Bell III,[1] B. C. Clark,[11] A. Wang,[2] T. J. McCoy,[12] M. E. Schmidt,[12] P. A. de Souza Jr.[13]

Mineral deposits on the martian surface can elucidate ancient environmental conditions on the planet. Opaline silica deposits (as much as 91 weight percent SiO_2) have been found in association with volcanic materials by the Mars rover Spirit. The deposits are present both as light-toned soils and as bedrock. We interpret these materials to have formed under hydrothermal conditions and therefore to be strong indicators of a former aqueous environment. This discovery is important for understanding the past habitability of Mars because hydrothermal environments on Earth support thriving microbial ecosystems.

Opaline silica deposits are an indicator of past aqueous activity. Some regions of Mars exhibit a thermal infrared spectral signature that has been interpreted to result from coatings of amorphous silica on rocks (*1*, *2*), although this interpretation is not unique (*3*). Results from the Mars rover Opportunity have suggested that opaline silica could be present in sulfate-rich sedimentary rocks at Meridiani Planum (*4*). The rover Spirit recently investigated the Eastern Valley between Home Plate and the Mitcheltree/Low Ridge complex (Fig. 1) in Gusev crater. Here we describe the discovery of silica-rich deposits in the Eastern Valley and farther east near sulfate-rich soil deposits.

Home Plate consists of laminated–to–cross-bedded tephra that shows evidence for a volcanic explosive origin, including a bomb sag produced when an ejected ~4-cm clast fell into deformable ash deposits (*5*). Mitcheltree Ridge and Low Ridge, located east of Home Plate, are partially eroded synclinal structures that expose tephra deposits (including lapillistones) capped by a deposit of vesicular basalt boulders. Soils in the Inner Basin ~250 m to the north (Samra) and ~50 m to the east (Tyrone) (Figs. 1 and 2) of Home Plate contain hydrated ferric sulfate deposits (*6*, *7*). The mobility of ferric iron under apparently oxidizing conditions, leading to ferric sulfates and oxides, is suggestive of low pH con-

ditions (*8*–*10*), and both deposits are interpreted to have originated under hydrothermal, acid-sulfate conditions (*6*). Although hydrated ferric sulfates dominate these deposits, Alpha Particle X-ray Spectrometer (APXS) (*11*) data also indicate that the nonsulfate component is enriched in SiO_2 relative to nearby rocks and soils (*6*), and Miniature Thermal Emission Spectrometer (Mini-TES) (*12*) spectra show an absorption feature at ~475 cm^{-1} that may be associated with amorphous silica.

The floor of the Eastern Valley exposes buff-colored, platy outcrops that we interpret on the basis of their composition, mineralogy, and fine-scale texture to be volcanic ash deposits altered under oxidizing acid-sulfate–dominated conditions. These outcrops are covered in places by light-toned clasts and are found in close association with outcrops that show a distinctive nodular appearance (Fig. 2).

Mini-TES thermal emission spectra acquired for an outcrop (Tyrone nodules in Fig. 1) near the Tyrone light-toned soil deposit and for light-toned nodular outcrops within the Eastern Valley (Fig. 2) exhibit the spectral signature associated with high silica contents (Fig. 3A). Although the outcrops are partially covered by windblown soil, the spectra contain a distinctive vibrational absorption feature at ~475 cm^{-1}, which is interpreted to result from the Si-O bending mode, and another at ~1110 cm^{-1}, which is interpreted to be the Si-O stretching mode of opaline silica (*13*). The spectra are distinct from those of quartz or cristobalite (*13*). Correcting for soil and dust contamination enhances the depths of these features (Fig. 3B).

The spectra include a prominent absorption near ~1260 cm^{-1}. This feature is atypical of laboratory spectra of opaline silica, which usually display only a shoulder in this location. Previous work has demonstrated that the depth of this feature is a function of the incidence angle, because the longitudinal optic mode dominates the transverse optic mode as the incidence angle increases (*14*). We confirmed this behavior with laboratory emission measurements for both synthetic amorphous silica and natural opaline silica samples (fig. S1). With increasing emission angle (where 0° is normal to the surface), the feature transitions from a shoulder to a prominent emissivity minimum. Because Mini-TES observations are obtained at emission angles ranging from 40° to nearly 90°, the spectra of opaline silica should display this distinctive feature.

During a drive on martian solar day (or "sol") 1148 (*15*), the rover's inoperative right front wheel (*16*) excavated light-toned soil deposits (fig. S2). Mini-TES spectra for these deposits (Fig. 3A) show the spectral features associated with opaline silica, including a weak emissivity maximum near 1625 cm^{-1} that results from the bending mode of bound water (*17*). In addition, a relatively broad and deep emissivity minimum centered at ~950 cm^{-1} is present and is interpreted as a transparency feature that arises in fine-grained particulate material where multiple scattering accentuates an otherwise weak absorption (*18*). Panoramic Camera (Pancam) (*19*) near-infrared reflectance spectra of the light-toned, high-Si soil deposits and nodules also reveal a distinct absorption feature that is consistent with the presence of H_2O in hydrated minerals (*20*).

We used the Microscopic Imager (MI) (*21*), APXS, and Mössbauer Spectrometer (MB) (*22*) to investigate several rocks and soils in the Eastern Valley. One rock outcrop [Elizabeth Mahon (*23*), which coincides approximately with the

[1]Department of Astronomy, Space Sciences Building, Cornell University, Ithaca, NY 14853, USA. [2]Department of Earth and Planetary Sciences, Washington University, St. Louis, MO 63130, USA. [3]School of Earth and Space Exploration, Arizona State University, Tempe, AZ 85287, USA. [4]Department of Physics, University of Guelph, Guelph, Ontario, N1G 2W1, Canada. [5]NASA Johnson Space Center, Houston, TX 77058, USA. [6]New Mexico Museum of Natural History and Science, Albuquerque, NM 87104, USA. [7]NASA Ames Research Center, Moffett Field, CA 94035, USA. [8]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA. [9]Department of Geosciences, State University of New York, Stony Brook, NY 11794, USA. [10]Department of Geological Sciences, University of Nevada, Reno, NV 89557, USA. [11]Lockheed Martin Corporation, Littleton, CO 80127, USA. [12]Department of Mineral Sciences, National Museum of Natural History, Smithsonian Institution, Washington, DC 20560, USA. [13]Tasmanian Information and Communication Technology Centre, Commonwealth Scientific and Industrial Research Organisation, Castray Esplanade, Hobart, TAS 7000, Australia.

*To whom correspondence should be addressed. E-mail: squyres@astro.cornell.edu

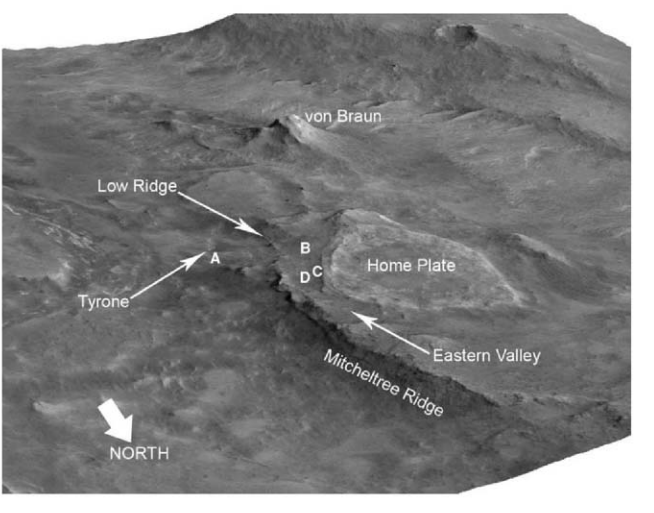


Fig. 1. Perspective view of the Inner Basin of the Columbia Hills. The Eastern Valley is the location of the silica-rich deposits, with the exception of an outcrop near Tyrone [**(A)**, Tyrone nodules outcrop]. **(B)** Location of the Kobal outcrop (Fig. 2A), **(C)** location of the Elizabeth Mahon outcrop, and **(D)** location of Nancy Warren, Innocent Bystander, and Gertrude Weise (Fig. 2B). The distance between the southern end of Home Plate and von Braun Hill is ~120 m. Image was generated using the Mars Reconnaissance Orbiter High Resolution Imaging Science Experiment (HiRISE) image PSP_001777_1655_red and a digital elevation model produced from HiRISE data.

Clara_Zaph4 Mini-TES target] had 72.4 weight percent (wt %) SiO_2 (table S1). Light-toned soils exposed by the rover's right front wheel (Kenosha Comets and Lefty Ganote) (Fig. 2B and fig. S2) had SiO_2 concentrations of 90.1 and 74.6 wt %, respectively. Two rock samples (Innocent Bystander and Norma Luker) (Fig. 2B and fig. S3) with, respectively, 63.1 and 69.2 wt % SiO_2 were rock fragments several centimeters in size that were created by breaking an outcrop with the rover's right front wheel. The high-silica measurements were made on the freshly broken surfaces, demonstrating that the silica is present within the rock to a depth of at least several centimeters. Microscopic Imager images of all outcrop targets show varying degrees of soil contamination, so these SiO_2 concentrations should be regarded as lower limits.

The APXS sampling depth is ~5 to 10 μm in rock for low–atomic number elements such as Si and 50 to 100 μm for higher–atomic number elements (*11*). This depth dependence can be used to investigate whether high-Si targets have a thin cover of soil. Compared with typical soils at the Spirit site, the light-toned soils are enriched in Si, Ti, Cr, and Zn (Fig. 4). Most other major elements, however, follow an exponential downward trend in relative abundance with rising element atomic number. This trend is a clear indicator that soil grains are mixed into the surface of the silica-rich material. On the basis of MB measurements and the Mg/Al ratio in APXS data, we estimate that soil contaminates ~30% of the Kenosha Comets target. (A small quantity of an alteration product could also be present.) Removing the contaminants leads to an inferred composition of ~98 wt % SiO_2, ~1.5 wt % TiO_2, ~0.4 wt % Cr_2O_3, ~200 parts per million (ppm) Ni, and ~330 ppm Zn.

APXS data also show that this soil target is rich in Si to a depth of at least 300 μm by comparing the ratios of the Compton and Rayleigh scatter peaks of the primary excitation radiation for x-ray fluorescence. The cross sections responsible for these two peaks have different atomic-number dependencies (including a large dependency on oxygen), and they originate from depths of as much as 300 μm. The scatter peaks ratio is nearly a factor of 2 different than it is for typical soils but is consistent with the scatter peak of a pure SiO_2 sample measured by the flight instrument during APXS instrument calibration (fig. S4).

Further evidence that the material is rich in SiO_2 to at least a few hundred micrometers is provided by nearly featureless Mössbauer spectra. The MB instrument samples to a few hundred micrometers (*22*) and only detects minerals that contain Fe. The material is therefore Fe-poor over at least this depth range.

The light-toned soils were mixed by the action of the rover wheels to a depth of a few centimeters, so the APXS and MB results imply that the soil is silica-rich to at least this depth. Additionally, because MI images show that they are dominated by grains much smaller than a few hundred micrometers in size, the APXS and MB data also demonstrate that these are not grains of typical martian soil coated with silica.

We interpret the silica-rich materials identified by Spirit to have formed under hydrothermal conditions, because they are found in close association with volcanic materials [e.g., Home Plate (*5*)] and, in some cases, are intimately mixed with ferric sulfates that are also probably of hydrothermal origin (*6*). Origin in a lacustrine evaporitic setting is unlikely given the complete lack of evidence for standing or flowing surface water at the Spirit site (*24*).

On Earth, a range of processes can produce high concentrations of opaline silica in hydrothermal settings. These processes result from the high solubility of silica in aqueous systems under warm and alkaline conditions (*25*) but much lower solubility at low pH. At one extreme, abundant hot neutral-to-alkaline groundwaters dissolve silica from subsurface rocks and then precipitate it when they reach the surface and cool and evaporate, forming sinter deposits. At the other extreme, very low pH waters are formed by condensation of water vapor and acid gases, typically in volcanic settings and at low water/rock ratios. These acid condensates interact with precursor rock like basalt, leaching away most minerals but leaving the most insoluble constituents—notably silica—behind (*26*). As an intermediate case, hydrothermal sinter deposits can also form at pH 2 to 4 and high water/rock ratios (*27*).

The enrichment of Ti (and Cr) in the silica-rich materials discovered by Spirit may provide a constraint on their origin. Like silica, titanium dioxide is relatively insoluble at low pH. A straightforward explanation for our observations therefore is that acid-sulfate low pH waters dissolved basaltic precursor materials, concentrating the highly insoluble silica (probably as opal-A) and titanium dioxide (probably as anatase). Based on the composition of relatively unaltered basalts at Gusev,

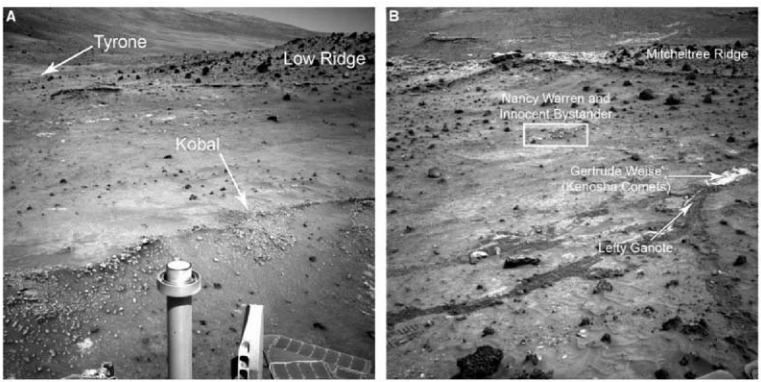


Fig. 2. (A) Navcam image acquired within the Eastern Valley, looking to the southeast and showing the nodular outcrop target Kobal that was examined with Mini-TES. The Tyrone area is shown before Spirit traversed there and exposed light-toned soils with its wheels. Navcam frame 2N195257205 acquired on sol 776. **(B)** Navcam image acquired in the Eastern Valley, looking west, showing the trench that was excavated by the front wheel and outcrops examined by the rover. Navcam frame 2N233253342 acquired on sol 1204.

titanomagnetite is the most likely precursor phase for the Ti.

Favoring a low pH origin is the similarity of silica-rich soils to other light-toned soil deposits within the Columbia Hills that are rich in hydrated sulfate salts, particularly ferric sulfates (*6*, *7*). Not only is their geologic setting similar—highly concentrated light-toned deposits excavated by the rover's wheels from just below the surface—but some of these sulfate-rich soils also contain a significant quantity (4 to 33%) of free silica (*6*). These observations may point toward a common origin, and as noted above, we interpret the hydrated ferric sulfate deposits to have formed under acid-sulfate conditions. Also, the compositional diversity of Si-rich materials observed by Spirit is consistent with formation by acid leaching of a variety of basaltic precursor materials, including those found in the Home Plate area (Fig. 5).

To remove nonvolatile components (e.g., Al, Fe, and Na) from the precursor rock, fluids—not vapors—are probably required. An unresolved issue is how far any fluids traveled before the most insoluble materials (SiO_2 and TiO_2) precipitated. If precipitation was local, at low water/rock ratios, the materials discovered by Spirit would most resemble the highly leached rocks and residual silica sometimes found in the vicinity of fumaroles (*26*, *28*). If there was considerable transport away from the source rocks before precipitation, implying higher water/rock ratios, the deposits would be more akin to the hydrothermal sinter deposits found at hot springs.

The composition of the Si-rich deposits does not uniquely constrain the pH of the fluids involved in their formation. For example, at high pH it is possible to precipitate sinter deposits from alkali chloride brines. Such precipitates are high in silica, and the presence of chloride can allow for elevated Ti levels as $TiCl_4$ or $TiOCl_2$ and the coprecipitation of anatase (*29*, *30*). Given the ample evidence for acid-sulfate processes elsewhere at Gusev, however, formation at low pH seems likely.

Hydrothermal conditions that produced high-purity opaline silica deposits on Mars could have led to locally habitable environments. Fumarolic environments and hot springs on Earth create habitable conditions for microbial life by conveying thermal waters, solutes, and volcanic gases to surface and near-surface environments (*31*, *32*) under both alkaline (*33*) and acidic (*31*, *34*) conditions. Any process that involves reprecipitation of silica from fluids can also provide a mechanism for preserving evidence of microbes (*35*). Whereas the silica-rich materials examined by Spirit are too small to be seen from orbit by current sensors, recent Compact Reconnaissance Imaging Spectrometer for Mars (*36*) observations of other regions on Mars may suggest the presence of more widespread and accessible hydrated and hydroxylated silica deposits elsewhere (*37*). Silica-rich deposits on Mars may be strong candidates for examination by future landed missions and as returned samples.

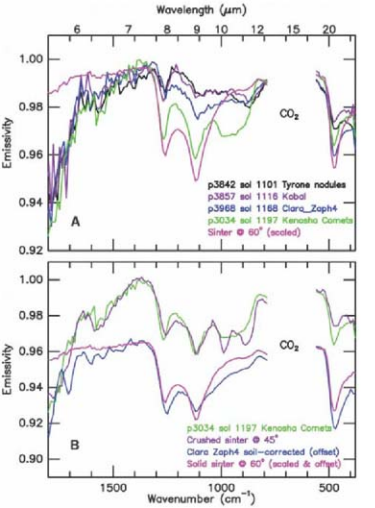


Fig. 3. (A) Mini-TES emissivity spectra for the outcrops Tyrone nodules, Kobal, Clara_Zaph4 (same location as the Elizabeth Mahon in situ target), and the soil Kenosha Comets. Also shown is a terrestrial opaline sinter acquired at an emission angle of 60° (similar to the Mini-TES observations). The notation pXXXX denotes the Mini-TES sequence identifier. **(B)** Comparison of a Mini-TES spectrum of Kenosha Comets with crushed hydrothermal sinter and comparison of a Mini-TES spectrum of the outcrop target Clara_Zaph4 (after correction for soil contamination) with solid hydrothermal sinter. All sinter samples are from Steamboat Springs, Nevada.

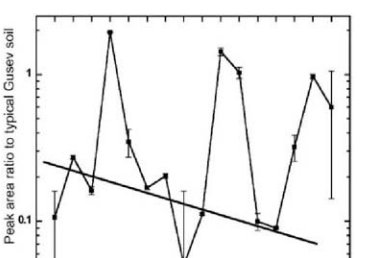


Fig. 4. Elemental composition of the bright soil deposit Kenosha Comets, ratioed on an element-by-element basis to typical Gusev soil. Si is strongly enriched relative to typical soil, and there are weaker enrichments in Ti, Cr, and Zn. Other major elements are strongly depleted relative to typical soil and exhibit an exponential decrease of relative abundance with increasing atomic number (diagonal line). Error bars indicate the 2σ statistical error in the Kenosha Comets sample.

References and Notes

1. M. D. Kraft *et al.*, *Geophys. Res. Lett.* **30**, 2288 (2003).
2. J. R. Michalski *et al.*, *Icarus* **174**, 161 (2005).
3. J. L. Bandfield, V. E. Hamilton, P. R. Christensen, *Science* **287**, 1626 (2000).
4. T. D. Glotch *et al.*, *J. Geophys. Res.* **111**, E12S06 (2006).
5. S. W. Squyres *et al.*, *Science* **316**, 738 (2007).
6. A. S. Yen *et al.*, *J. Geophys. Res.* **113**, E06S10 (2008).
7. J. R. Johnson *et al.*, *Geophys. Res. Lett.* **34**, L13202 (2007).
8. N. J. Tosca, A. Smirnov, S. M. McLennan, *Geochim. Cosmochim. Acta* **71**, 2620 (2007).
9. G. M. Marion, J. S. Kargel, D. C. Catling, *Geochim. Cosmochim. Acta* **72**, 242 (2008).
10. N. J. Tosca *et al.*, *J. Geophys. Res.*, 10.1029/2007JE003019.
11. R. Rieder *et al.*, *J. Geophys. Res.* **108**, 8066 (2003).
12. P. R. Christensen *et al.*, *J. Geophys. Res.* **108**, 8064 (2003).
13. J. R. Michalski *et al.*, *Geophys. Res. Lett.* **30**, 2008 (2003).
14. R. M. Almeida, *Phys. Rev. B* **45**, 161 (1992).
15. A martian solar day has a mean period of 24 hours, 39 min, and 35.244 s and is referred to as a "sol" to distinguish it from the shorter solar day on Earth. Spirit crossed from the Gusev plains to the Columbia Hills on the 157th sol after landing. Observations discussed in this paper were made over the period from sols 749 to 1288.
16. Spirit's right front wheel drive actuator failed on sol 799. Since then, the rover has mainly driven backward

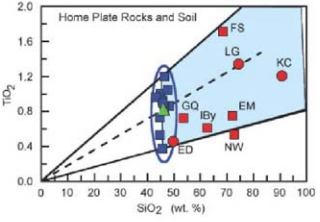


Fig. 5. Plot of titanium dioxide and silica contents for APXS measurements acquired since arriving at Home Plate and including observations from the Eastern Valley, Low Ridge, Mitcheltree Ridge, and the Tyrone areas. Red squares, silica-rich rocks; red circles, silica-rich soils; blue squares, basaltic rocks in the vicinity of Home Plate; and green triangle, typical local soil. The light blue region represents the compositions that can be obtained by acid-sulfate leaching of Home Plate rocks, assuming no variation in the TiO_2/SiO_2 ratio. KC, Kenosha Comets; LG, Lefty Ganote; FS, Fuzzy Smith; EM, Elizabeth Mahon; NW, Nancy Warren; ED, Eileen Dean; Iby, Innocent Bystander; and GQ, Good Question. The dashed line represents a typical evolutionary trend for leaching that preserves SiO_2 and TiO_2, and the blue oval indicates the range of basaltic rock compositions in the vicinity of Home Plate.

and dragged this nonrotating wheel along the surface, scraping paths through soils and crushing weak rocks.
17. J. W. Salisbury *et al.*, *Infrared (2.1-25 μm) Spectra of Minerals* (Johns Hopkins Univ. Press, Baltimore, 1991).
18. J. W. Salisbury *et al.*, *J. Geophys. Res.* **92**, 702 (1987).
19. J. F. Bell III *et al.*, *J. Geophys. Res.* **108**, 8063 (2003).
20. M. S. Rice *et al.*, Abstract #2138, presented at the 39th Annual Lunar and Planetary Science Conference, Houston, TX, 10 to 14 March 2008.
21. K. E. Herkenhoff *et al.*, *J. Geophys. Res.* **108**, 8065 (2003).
22. G. Klingelhöfer *et al.*, *J. Geophys. Res.* **108**, 8067 (2003).
23. Most of the rock and soil targets on the east side of Home Plate are named after players and teams from the All American Girls Professional Baseball League that operated from 1943 to 1954. The names are informal and not approved by the International Astronomical Union.
24. R. E. Arvidson *et al.*, *J. Geophys. Res.* **111**, E02S01 (2006).
25. P. M. Dove, J. D. Rimstidt, in *Silica: Physical Behavior, Geochemistry and Materials Applications*, P. J. Heaney, C. T. Prewitt, G. V. Gibbs, Eds. (Mineralogical Society of America, Washington, DC, vol. 29, 1994) pp. 259–308.
26. R. V. Morris *et al.*, *Lunar Planet. Sci. XXXI*, 2014 (2000).
27. K. A. Rodgers *et al.*, *Earth Sci. Rev.* **66**, 1 (2004).
28. K. A. Rodgers *et al.*, *Clay Miner.* **37**, 299 (2002).
29. H. D. Nam *et al.*, *Jpn. J. Appl. Phys.* **37**, 4603 (1998).
30. S. J. Kim, S. D. Park, Y. H. Jeong, *J. Am. Ceram. Soc.* **82**, 927 (1999).
31. W. P. Inskeep, T. R. McDermott, *Geothermal Biology and Geochemistry in Yellowstone National Park* (Thermal Biology Institute, Bozeman, MT, 2003).
32. M. R. Walter, D. J. Des Marais, *Icarus* **101**, 129 (1993).
33. W. N. Doemel, T. D. Brock, *Appl. Environ. Microbiol.* **34**, 433 (1977).
34. P. L. Slering *et al.*, *Geomicrobiol. J.* **23**, 129 (2006).
35. S. L. Cady, J. D. Farmer, in *Evolution of Hydrothermal Ecosystems on Earth (and Mars?)*, G. Bock, J. Goode, Eds. (J. Wiley and Sons, Chichester, UK, 1996), pp. 150–173.
36. S. Murchie *et al.*, *J. Geophys. Res.* **112**, E05S03 (2007).
37. R. E. Milliken *et al.*, Abstract P22A-02, presented at the American Geophysical Union Fall Meeting, San Francisco, CA, 10 to 14 December 2007.
38. This research was carried out for the Jet Propulsion Laboratory, California Institute of Technology, under a contract with NASA.

Supporting Online Material

www.sciencemag.org/cgi/content/full/320/5879/1063/DC1
Figs. S1 to S4
Table S1

18 January 2008; accepted 15 April 2008
10.1126/science.1155429

Formation of Box Canyon, Idaho, by Megaflood: Implications for Seepage Erosion on Earth and Mars

Michael P. Lamb,* William E. Dietrich, Sarah M. Aciego, Donald J. DePaolo, Michael Manga

Amphitheater-headed canyons have been used as diagnostic indicators of erosion by groundwater seepage, which has important implications for landscape evolution on Earth and astrobiology on Mars. Of perhaps any canyon studied, Box Canyon, Idaho, most strongly meets the proposed morphologic criteria for groundwater sapping because it is incised into a basaltic plain with no drainage network upstream, and approximately 10 cubic meters per second of seepage emanates from its vertical headwall. However, sediment transport constraints, ^{4}He and ^{36}Cl dates, plunge pools, and scoured rock indicate that a megaflood (greater than 220 cubic meters per second) carved the canyon about 45,000 years ago. These results add to a growing recognition of Quaternary catastrophic flooding in the American northwest, and may imply that similar features on Mars also formed by floods rather than seepage erosion.

A central thrust in geomorphology and planetary science is to link diagnostic landscape morphologies to formation processes. A prominent example is the formation of amphitheater-headed canyons, in which the stubby appearance of valley heads, steep headwalls, and little landscape dissection upstream have long been interpreted to result from seepage erosion or groundwater sapping on Earth (*1*–*4*), Mars (*5*, *6*), and now Titan (*7*). Theory (*8*), experiments (*9*), and field studies (*10*) have validated this hypothesis in unconsolidated sand, showing that valley heads are undermined and propagate upstream from seepage-induced erosion. This means that valleys can grow without precipitation-fed overland flow, which has profound implications for landscape evolution on Earth and the hydrologic cycle and habitability of Mars.

Despite widespread acceptance of the seepage-erosion hypothesis and its validation in sand, we lack an unambiguous example of an amphitheater-headed canyon formed by seepage erosion in bedrock because of overlapping features generated by rainfall runoff at most sites on Earth (*11*). Even the amphitheater-headed valleys of the Colorado Plateau and Hawaii, which are most often cited as classic examples of groundwater sapping in bedrock (*2*, *3*), have been in question because of evidence for flash floods and plunge-pool erosion (*11*–*13*). To better evaluate the seepage-erosion hypothesis, we set out to study the erosion and transport processes within a bedrock canyon, Box Canyon, Idaho, USA, which has a steep amphitheater-shaped headwall, contains the 11th-largest spring in the United States, and lacks the landscape dissection and rainfall runoff upstream of its headwall that has made other sites controversial (Fig. 1A). Moreover, Box Canyon exhibits remarkable similarity in morphology and possibly lithology (basalt) with many Martian canyons (Fig. 1B) that have been attributed to seepage erosion (*5*, *6*).

Box Canyon is located within the Snake River Plain, a broad and relatively flat basin in southern Idaho filled by sediments and volcanic flows that erupted over the course of ~15 million to 2 thousand years ago (ka) (*14*). Several tributaries of the Snake River Canyon appear as stubby valleys that end abruptly in amphitheater heads, including Malad Gorge, Blind Canyon, and Box Canyon (Fig. 2), all of which have been attributed to seepage erosion (*1*, *4*). Box Canyon is cut into the Sand Springs Basalt [also named the Basalt of Rocky Butte (*15*); see supporting online material (SOM) text] with an Ar-Ar eruption age of 95 ± 10 ka (*16*) and U-Th/He eruption ages that range from 86 +12 ka to 130 ± 12 ka (*17*), and this basalt filled an ancestral canyon of the Snake River (*18*).

Department of Earth and Planetary Science, University of California, Berkeley, CA 94720–4768, USA.

*To whom correspondence should be addressed. E-mail: mpl@berkeley.edu

The permeable lava beds of the Snake River Plain form an extensive aquifer, with recharge entering in the east [for example, Big Lost River sinks (Fig. 2)] and groundwater flowing westward. Large springs emanate from the east wall of the Snake River Canyon between Box Canyon and Malad Gorge, where the river jogs north, perpendicular to the regional topographic slope and the groundwater flow direction. These springs have a cumulative discharge of ~170 m^3/s, and one of the largest [~10 m^3/s (fig. S6)] emanates from the head of Box Canyon, creating Box Canyon Creek (*19*).

Box Canyon is sinuous (Fig. 3A), and the longitudinal profile is approximately 2.68 km in length and has an average channel-bed slope of 2.18% (Fig. 3B and fig. S3). The canyon is ~35 m deep and 120 m wide at its head and about twice as deep and wide at its mouth. The columnar basalt walls of the canyon have collapsed, creating steep (~20° to 35°) talus slopes, many of which abut Box Canyon Creek. Talus accumulation lessens upstream and is absent at the canyon head (Fig. 4A). Several terracelike platforms are elevated 2 to 7 m above the current stream level and separate the steep talus slopes from the creek (Fig. 3 and fig. S3). These contain large boulders (>1 m), and some appear to be imbricated in the downstream direction, indicating past fluvial transport.

Although the lack of talus at the canyon head has been taken as evidence of continued dissolution (*4*), the bedrock composing the headwall and surrounding talus are blocky and hard and show no visual evidence for enhanced weathering. Water samples from Box Canyon Creek and neighboring wells have silica concentrations of 32 to 35 mg/l, which bracket the saturation value (see SOM text and fig. S6), suggesting that the groundwater is in equilibrium with the basaltic aquifer and that substantial dissolution is not occurring at Box Canyon spring.

Despite no modern overland flow contribution to Box Canyon Creek, three features at the canyon head indicate that surface water once flowed into the canyon. First, three concentric semicircles of boulders within the canyon head appear to be waterfall plunge pools with ~2 m of relief (Fig. 4A). Second, a small notch (~300 m^3) in the center of the headwall rim (Fig. 4A) has linear flutelike abrasion marks, millimeters in width and several centimeters long, that follow the local curvature of the notch, indicating past overspill. The scours appear as divots on the inferred upstream end that gradually fan outward and diminish in relief downstream (Fig. 4B). Third, this scoured rock extends at least 1 km upstream of the canyon head and delineates flow toward the canyon (Fig. 3). The scoured path cannot be followed further upstream because it is covered by loess, deposited from ~40 to 10 ka (*20*).

The basalt in Box Canyon breaks down into large boulders (~1 m) that, without dissolution, must be transported downstream to allow canyon growth. Despite the great discharge of the spring, no measurable amount of sediment is currently transported. A minimum estimate of flow needed to carve the canyon can be found by calculating the discharge necessary to initiate sediment transport on the creek bed. We measured channel cross sections (fig. S4), longitudinal channel-bed profiles (Fig. 3 and fig. S3), and grain-diameter distributions (fig. S2) [diameter (D_{84}) = 0.6 m,

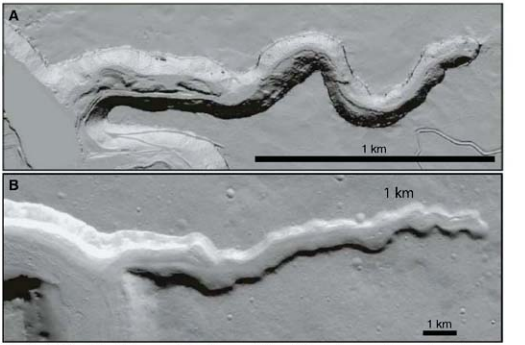


Fig. 1. (A) Shaded relief map of Box Canyon, Idaho. Airborne laser-swath mapping data were collected by the National Center for Airborne Laser Mapping. The data have been filtered to remove vegetation that exists along the creek banks. This is a Universal Transverse Mercator (UTM) zone 11 projection, North American Datum of 1983 (NAD83) datum, at 1 m resolution. **(B)** Thermal Emission Imaging System (*32*) infrared daytime image of Mamers Vallis, Mars, image V19470014, at 19 m resolution.

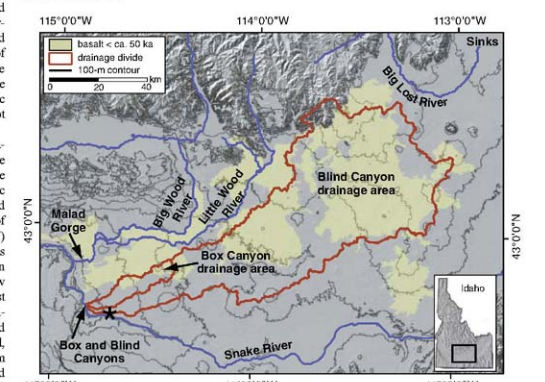


Fig. 2. Topographic map of the eastern Snake River Plain, the location of which is shown on the inset map of Idaho. The drainage areas feeding Box Canyon (228 km^2) and Blind Canyon (4713 km^2) are outlined, following the path of steepest descent. The yellow-shaded regions mark the locations of volcanism younger than ~50 ka (*15*). The thin black lines are 100-m topographic contours. Topographic data are from the U.S. Geological Survey. This is a UTM zone 11 projection, NAD83 datum, at 25 m resolution.

D_{50} = 0.29 m, and D_{16} = 0.13 m; the subscripts denote the percentage of grains that are finer] at a relatively straight 125-m reach within the canyon (Fig. 3A). A critical Shields stress formula for incipient motion (*21*) combined with our measured channel cross section (XS2 in Fig. 3A and fig. S4), local average bed slope (S = 1.85%; Fig. 3B and fig. S3), and a flow-resistance equation (*22*) determined that a flow discharge $Q > 220$ m^3/s (corresponding to an average flow depth $h > 1.7$ m) would be necessary to move the sediment bed and continue canyon erosion (*23*). This is a factor of 22 greater than the modern spring discharge ($Q \sim 10$ m^3/s).

The scoured-rock upslope of the canyon head occurs within a broad channel-like depression ~250 m wide and 3 m deep (XS1 in Fig. 3A and fig. S4). The scours extend over the southern bank of XS1, indicating that flow was deeper than and only partially bounded by this channel. A discharge estimate can be made for the flood event that spilled over the canyon rim by assuming that the flow was contained within this channel. Using the measured cross-sectional area at the threshold of overspill at XS1 (475 m^2), the regional bedrock slope parallel to scour marks (S = 0.74%), a flow-resistance formula (*22*), and a wide range in of bed-roughness length scales $0.1 \leq k_s \leq 1$ m (because this is the least-constrained parameter), we calculated a minimum flow discharge ranging from 800 to 2800 m^3/s (*23*). This would have filled the canyon to a depth of 3.7 to 5.8 m within our measurement reach (fig. S4), and, unlike seepage, would have exceeded the competency threshold to transport the bouldery bed. These estimated discharges are large, but smaller than the peak discharge of other catastrophic floods in the region [for example, the Bonneville flood, 10^6 m^3/s (*24*), and the Big Lost River Flood, 6×10^4 m^3/s (*25*)].

The vertical profile of the headwall suggests that it migrated upstream as a knickpoint, and the near-vertical joints inherent to flood basalt probably promoted toppling of basalt columns. The lack of gravel upstream of the canyon head also limited abrasion of the canyon rim. If sediment transport was the rate-limiting step for canyon erosion, a duration of flow needed to carve the canyon can be estimated by dividing the total volume of the canyon (~1.53 × 10^7 m^3) by a volumetric transport rate of sediment (*26*) for our estimates of flood discharge (800 to 2800 m^3/s). This suggests that flow was sustained for 35 to 160 days to transport the required load out of the canyon (*23*), which is similar to the estimated duration of the Bonneville flood [~100 days (*24*)]. Excavation of Box Canyon could have taken less time, however, because the flood was only partially contained within the channel at XS1.

We collected four samples, distributed in the streamwise direction within the canyon (Fig. 3), for ^{3}He cosmogenic-exposure–age dating to further constrain the history of the canyon. We sampled scoured bedrock that was exposed at the canyon-head rim [location 4 (Fig. 4B)], and three large boulders that, from their size and separation from the active talus slopes, appeared to be relatively stable (fig. S1B). Active talus production from canyon walls, as well as weathering, means that the ages for these boulder surfaces provide minimum ages of the canyon.

Of the boulders sampled, only location 2 was on a terrace among other large imbricated boulders, potentially indicating past fluvial transport (Fig. 3 and fig. S1B). This sample yielded an exposure age of 48 ± 3 ka (1 σ error) and the other two boulders were nearly half as old

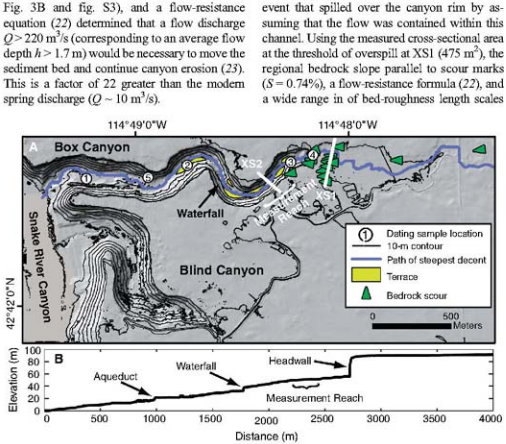


Fig. 3. **(A)** High-resolution topographic map of Box and Blind Canyons. The yellow-shaded regions mark potential fluvial terraces, which range from 2 to 7 m above the current stream level (fig. S3). Detailed measurements of flow depth, water surface slope, bed slope, channel width, and bed particle size were made within the region marked "measurement reach" (*23*). Discharge calculations were made using cross-sectional areas measured at XS1 and XS2 (fig. S4). Mapped scours on bedrock (Fig. 4B) are shown as green arrows. The white circles are sample locations used for dating. The thin black lines are 10-m topographic contours. The blue line is the calculated path of steepest descent, but does not indicate modern-day flow paths because no flow on record has spilled over the canyon headwall. See Fig. 1 for data source and projection. **(B)** Longitudinal profile of Box Canyon extracted from light detection and ranging data (Fig. 3A), following the path of steepest descent. Major breaks in slope correspond to the canyon headwall, waterfall, and a disturbed region near the canyon mouth caused by an aqueduct.

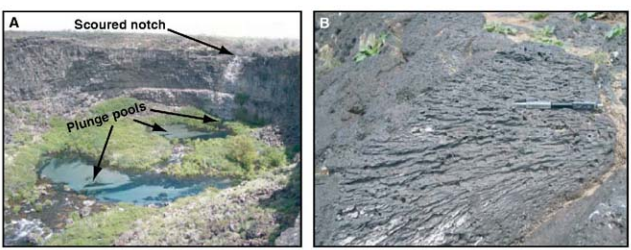


Fig. 4. **(A)** Photograph of the head of Box Canyon. The three concentric circles that lack boulders are interpreted to be plunge pools. The headwall relief is ~35 m. **(B)** Photograph of scours within the notch of the Box Canyon headwall. The pencil included for scale is ~14 cm long and points in the inferred flow direction.

(location 1, 21 ± 1 ka; location 3, 19 ± 3 ka) (*17*, *23*). The scoured notch (location 4) yielded the same age as the boulder at location 2, 45 ± 5 ka.

Another constraint on the age of the canyon comes from an ~20-cm-thick, finely laminated bed, containing clay, silt, and sand, that is exposed in a small road cut within the talus slope [location 5 (Fig. 3 and fig. S1C)]. Two shells found within the layer yielded ages of 22.4 ± 1 × 10^3 radiocarbon years before the present (*23*), which is equivalent to a calibrated age of ~26 ka (*27*).

Together, these observations, hydraulic calculations, and dates eliminate the seepage-erosion hypothesis for the formation of Box Canyon. Here, seepage is not substantially enhancing the weathering of the headwall, and contemporary seepage flow is deficient by a factor of ~22 to evacuate sediment from the canyon. Moreover, erosion of the canyon headwall ceased ~45 ka according to the age of the scoured-bedrock notch. The observations of scoured bedrock and plunge pools point toward a flood or floods able to transport boulders and carve the canyon in weeks or months, illustrating the power of rare catastrophic events in shaping landscape. This hypothesis is supported by the similarity in surface-exposure ages of the terrace-bound boulder at location 2 and the scoured notch. We interpret the younger boulders to have rolled to the canyon floor after the canyon was formed, which is consistent with the shell deposit sandwiched between several meters of talus (fig. S1C), indicating that canyon formation occurred well before 22 ka and wall collapse has been active since. Whereas a single flood event is the simplest interpretation, we cannot rule out multiple events occurring after 86 ± 12 ka [the eruption age of the basalt (*17*)], with the last event resetting the exposure ages to ~45 ka.

The dates indicate that the canyon is much older than the Bonneville flood that occurred within the Snake River Canyon ~14.5 ka. Given the drainage area of Box Canyon (228 km^2) and Blind Canyon (4713 km^2) (Fig. 2), and our flood-duration and discharge estimates, a sufficient meteorological flood would require more than 1.7 m of runoff lasting for several weeks or longer, which is highly unlikely because modern annual precipitation averages only 0.22 m over the eastern Snake River Plain, infiltration is extremely high, and conditions were probably drier ~45 ka (*28*). Potential diversions of the Snake River by lava flows also seem unlikely because the Snake River Canyon and the Box and Blind drainages are separated topographically, except for one location (marked with an asterisk on Fig. 2), and no volcanic dams with an age of ~45 ka have been discovered there.

The most likely source for the Box Canyon flood is the Little and Big Wood River drainage basins to the north or the Big Lost River drainage basin to the northeast (Fig. 2). Both drainages produced large-magnitude paleofloods that cut canyons in Quaternary basalt, scoured bedrock, and transported large (~1 m) boulders (*15*, *25*). For example, the paleomegaflood of the Big Lost River, which occurred sometime between ~19 ka and 95 ka (*29*), was probably an outburst from the Pleistocene Glacial-Lake East Fork and had a peak flow of 6 × 10^4 m^3/s (*25*). Such an event would have easily surpassed the drainage divide between Box Canyon and the Wood and Lost River drainages (Fig. 1). The divides themselves also have shifted since the formation of Box Canyon because of volcanism that postdates the Box Canyon flood (Fig. 2).

Our findings suggest that Box Canyon and other amphitheater-headed canyons carved into basalt by large floods [for example, Dry Falls, Washington, USA (*30*), and Ásbyrgi Canyon, Iceland (*31*)] might be better terrestrial analogs of Martian canyons in volcanic terrains than seepage channels in sand.

References and Notes

1. I. C. Russel, *U.S. Geol. Surv. Bull.* **199**, 192 (1902).
2. J. E. Lahy, M. C. Malin, *Geol. Soc. Am. Bull.* **96**, 203 (1985).
3. R. C. Kochel, J. F. Piper, *J. Geophys. Res.* **91**, E175 (1986).
4. H. T. Stearns, *J. Geol.* **44**, 429 (1936).
5. M. C. Malin, M. H. Carr, *Nature* **397**, 589 (1999).
6. K. P. Harrison, R. E. Grimm, *J. Geophys. Res.* 10.1029/2003JE002455 (2005).
7. M. G. Tomasko *et al.*, *Nature* **438**, 10.1038/nature04126 (2005).
8. T. Dunne, *Prog. Phys. Geogr.* **4**, 211 (1980).
9. A. D. Howard, C. F. McLane, *Water Resour. Res.* **24**, 1659 (1988).
10. S. A. Schumm, K. F. Boyd, C. G. Wolff, W. J. Spitz, *Geomorphology* **12**, 281 (1995).
11. M. P. Lamb *et al.*, *J. Geophys. Res.* **111**, 10.1029/2005JE002663 (2006).
12. M. P. Lamb, A. D. Howard, W. E. Dietrich, J. T. Perron, *Geol. Soc. Am. Bull.* **119**, 805 10.1130/B25986.1 (2007).
13. A. D. Howard, W. E. Dietrich, M. A. Seidl, *J. Geophys. Res.* **99**, 13971 (1994).
14. H. E. Malde, in *Quaternary Nonglacial Geology: Conterminous U.S.*, R. B. Morrison, ed. (Geological Society of America (GSA), Boulder, CO, 1991), vol. K-2.
15. J. D. Kauffman, K. L. Otherberg, V. S. Gillerman, D. L. Garwood, *Geologic Map of the Twin Falls 30 × 60 minute Quadrangle, Idaho* (Idaho Geological Survey, Moscow, ID, 2005).
16. L. Taure, C. Luskin, P. Selkin, P. Gans, A. Calvert, *Geochem. Geophys. Geosyst.* **5**, 10.1029/2003GC000661 (2004).
17. S. M. Aciego *et al.*, *Earth Planet. Sci. Lett.* **254**, 288 10.1016/j.epsl.2006.11.039 (2007).
18. H. E. Malde, *U.S. Geol. Surv. Prof. Pap.* **20**, 20 (1971).
19. U.S. Geological Survey: Box Canyon creek, gauge 13095500. This well log information can be found at http://waterdata.usgs.gov/id/nwis.
20. S. L. Forman, R. P. Smith, W. R. Hackett, J. A. Tullis, P. A. McDaniel, *Quat. Res.* **40**, 30 (1993).
21. M. P. Lamb, W. E. Dietrich, J. Venditti, *J. Geophys. Res.*, doi:10.1029/2007JF000831 (2008).
22. J. C. Bathurst, *J. Hydrol. (Amst.)* **269**, 11 (2002).
23. Materials and methods are available as supporting material on *Science* Online
24. J. E. O'Connor, *Hydrology, Hydraulics and Geomorphology of the Bonneville Flood*, (GSA Special Paper 274, GSA, Boulder, CO, 1993), p. 90.
25. S. L. Rathburn, *Geomorphology* **8**, 305 (1993).
26. R. Fernandez Luque, R. van Beek, *J. Hydraul. Res.* **14**, 127 (1976).
27. P. J. Reimer *et al.*, *Radiocarbon* **46**, 1029 (2004).
28. D. B. Madsen *et al.*, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **167**, 243 (2001).
29. T. E. Cerling, R. J. Poreda, S. L. Rathburn, *Geology* **22**, 227 (1994).
30. J. H. Bretz, *J. Geol.* **31**, 617 (1923).
31. H. Tomasson, *Natturufraeoingurinn* **43**, 12 (1973).
32. P. R. Christensen *et al.*, *Space Sci. Rev.* **110**, 85 (2004).

33. Funding was provided by the NASA Astrobiology Institute and the National Center for Earth Surface Dynamics. We thank C. May, J. McKean, T. Perron, R. Thurow, and T. van Soest for field assistance.

Supporting Online Material

www.sciencemag.org/cgi/content/full/320/5879/1067/DC1
Materials and Methods
SOM Text
Figs. S1 to S6
Table S1

19 February 2008; accepted 21 April 2008
10.1126/science.1156630

Anticorrelated Seismic Velocity Anomalies from Post-Perovskite in the Lowermost Mantle

Alexander R. Hutko,[1]* Thorne Lay,[1]† Justin Revenaugh,[2] Edward J. Garnero[3]

Earth's lowermost mantle has thermal, chemical, and mineralogical complexities that require precise seismological characterization. Stacking, migration, and modeling of over 10,000 *P* and *S* waves that traverse the deep mantle under the Cocos plate resolve structures above the core-mantle boundary. A small −0.07 ± 0.15% decrease of *P* wave velocity (V_p) is accompanied by a 1.5 ± 0.5% increase in *S* wave velocity (V_s) near a depth of 2570 km. Bulk-sound velocity [$V_b = (V_p^2 - 4/3V_s^2)^{1/2}$] decreases by −1.0 ± 0.5% at this depth. Transition of the primary lower-mantle mineral, $(Mg_{1-x-y}\ Fe_xAl_y)(Si,Al)O_3$ perovskite, to denser post-perovskite is expected to have a negligible effect on the bulk modulus while increasing the shear modulus by ~6%, resulting in local anticorrelation of V_b and V_s anomalies; this behavior explains the data well.

Increasing pressure (*P*) and temperature (*T*) with depth in Earth causes minerals to undergo phase transitions to new crystalline structures accompanied by abrupt changes in density (ρ), isentropic bulk modulus (K_S), and shear modulus (*G*) that result in seismic *P* wave velocity $\{V_p = [(K_S + 4/3G)/\rho]^{1/2}\}$ and *S* wave velocity $[V_s = (G/\rho)^{1/2}]$ discontinuities. The pri-

mary mineral in the lower mantle, magnesium-perovskite [$(Mg_{1-x-y}Fe_xAl_y)(Si,Al)O_3$] (Pv), has recently been discovered to undergo a phase transition to a 1 to 1.5% denser polymorph called post-perovskite (pPv) in the lowermost mantle (*1*–*4*). This phase change is unusual in that it involves a substantial increase in G with little change in K_S. Theoretical estimates of high-temperature elasticity for the two phases of pure $MgSiO_3$ suggest that V_s increases by ~2% but V_p changes much less (~±0.5%) across the transition (*5*–*7*). The bulk-sound velocity [$V_\Phi = (V_p^2 - 4/3V_s^2)^{1/2} = (K_S/\rho)^{1/2}$] for pPv is thus anomalously low, by about 1%, as compared with that for Pv. The reliable seismic calculation of V_Φ structure requires precise V_s and V_p structures in the same location, which is viable in only a few places. Here we provide a direct test for the presence of pPv in the deep mantle by determining the V_Φ and V_s structures in a region having a strong V_s discontinuity and a positive V_s volumetric anomaly. Confirmation of the regional presence of pPv in the deep mantle allows mineral-physics estimates of its properties to be applied to the interpretation of the deep seismological observations.

We studied the lowermost mantle beneath the Cocos Plate, which is well sampled by *P* and *S* waves from South American earthquakes recorded by seismic stations in western North America (Fig. 1). This region is known to have higher-than-average V_s (*8*–*11*) and slightly higher-than-average V_p (*8*, *12*). Investigations of deep mantle structure below the Cocos region have consistently yielded models with 1 to 2% V_s discontinuities situated 200 to 300 km above the core-mantle boundary (CMB) (*13*–*16*), but a large range of V_p discontinuity models involve no detectable discontinuity (*16*), a small (~0.5%) increase (*17*), or a large (–1 to –3%) decrease (*18*).

To resolve the V_p structure under the Cocos region, we collected 17,550 high-quality *P* waves from 243 earthquakes in South American subduction zones recorded by >800 broadband and short-period seismic stations in and near California. These data provide dense sampling of structure near the CMB, with ray paths spanning epicentral distances from 60° to 80°. *P* wave reflections from the CMB (*PcP*) sample the same region (5° to 15°N) as *S* wave reflections from the CMB (*ScS*) (Fig. 1) for a high-quality *S* wave data set that has previously been analyzed (*11*, *13*, *14*). This allows us to determine precise V_p and V_s structures in the same localized region.

We imposed a demanding signal-quality criterion on the *P* waves to ensure stable down-going radiation for the *PcP* portion of the wavefield. At distances beyond 60°, *PcP* typically have amplitudes <20% of *P* wave (fig. S1). Combining data from many events requires the equalization of the signal shapes, which we achieved by deconvolving source wavelets obtained by alignment and linear stacking of the direct *P* wave arrivals for each event separately. We then filtered the deconvolved signals for each event in various frequency passbands and stacked the signals on the expected travel time for *PcP*; if a stable, high signal-to-noise–ratio *PcP* image was not formed for a given passband, the event was discarded. This process eliminated many data, including numerous events used in earlier V_Φ studies, having either poor down-going radiation patterns or excessive differences in the waveshapes of the direct *P* wave and *PcP*. The resulting data set has relatively uniform detections of *PcP* and is sufficiently large and high-quality that we can identify features that have amplitudes only 10% as large as the weak *PcP* reflections. Our screened data set includes more than 10,000 waveforms from 75 events (Fig. 1).

The *S* wave data were equalized by deconvolving source wavelets obtained from aligning and stacking the clear *ScS* arrivals; only records having impulsive deconvolved *ScS* signals were retained. The total number of traces, 255, is much smaller than that for the *P* waves, but the *S* wave reflections from the lower mantle are far stronger than their *P* wave counterparts. Good resolution of V_s structure was obtained, as is critical for computing V_Φ.

Double-array stacking (*11*, *17*, *19*) (fig. S2) was applied to the *P* and *S* wave data sets to develop localized one-dimensional (1D) velocity structures for the two subregions indicated in Fig. 1 (5° to 10°N and 10° to 15°N). Stacks for the 7300 *P* waves sampling the two subregions are shown in Fig. 2 for the passband with the highest signal quality (0.25 to 0.5 Hz); corresponding subregion stacks for the 200 *S* waves from (*20*) are shown in fig. S3 for the passband ranging from 0.01 to 0.3 Hz. Alignment of the *P* wave data on *PcP* reflections accounted for a systematic northward increase in lowermost-mantle V_p across the region (fig. S4), but the stacks are similar to those found when the data are aligned on direct *P* wave arrivals (fig. S5B). Subdivision of the data bins into smaller latitude intervals of 2.5° confirms the stability of the basic features in the bins (fig. S5, C and D). The low amplitude (<2% of the direct *P* wave) of any *P* wave reflectivity at depths above the *PcP* arrival directly indicates that, at most, only tiny *PcP* precursors come from the lowermost mantle in this region. In contrast, *ScS* precursors from the lowermost mantle that are 25% as large as direct *S* waves are observed in this region (fig. S3), as has long been known (*14*).

Modeling of the *P* wave data stacks (*19*) demonstrates that only small velocity variations are present and bounds the depths of abrupt changes to within ±10 km. We therefore constrained the depth of any associated V_s features to

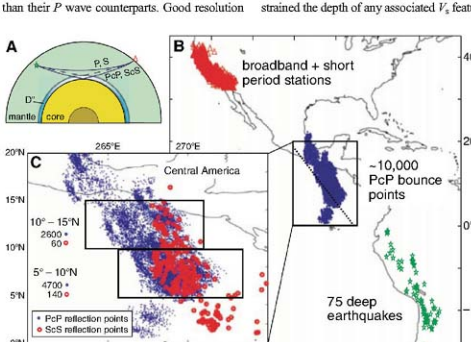


Fig. 1. (A) Earth cross section with representative ray paths for direct (*PiS*) and CMB reflected (*PcPiScS*) phases and any reflections from deep mantle discontinuities (dashed line). The D'' region structure in the lowermost mantle is the focus of this study. **(B)** Map indicating the study configuration, involving 75 earthquake epicenters (green stars), seismic stations (red triangles), and surface projections of *PcP* CMB reflection points (blue dots). The dotted line shows the surface trace of a cross section made through the migration image volume in the lower mantle (Fig. 3). **(C)** The inset map shows both *PcP* and *ScS* CMB reflection points and two data bins where these overlap, used in waveform stacking analysis.

[1]Earth and Planetary Sciences Department, University of California, Santa Cruz, CA 95064, USA. [2]Department of Geology and Geophysics, University of Minnesota, 310 Pillsbury Drive Southeast, Minneapolis, MN 55455, USA. [3]School of Earth and Space Exploration, Arizona State University, Tempe, AZ 85287–1404, USA.

*Present address. U.S. Geological Survey, Mail Stop 966 Box 25046, Denver Federal Center, Denver, CO 80225, USA. †To whom correspondence should be addressed. E-mail: thorne@pmc.ucsc.edu

match the V_P structures in modeling the S wave stacks, slightly modifying previous V_S models (*20*). We assumed that the largest amplitude features in the P and S wave data stacks near 300 km above the CMB represent reflectivity from colocated V_S and V_P discontinuities. This maximizes the reflection coefficient and provides upper bounds on any possible V_P discontinuity compatible with the weak P wave reflectivity near this depth.

For the 10° to 15° bin, our preferred model has a seismic velocity discontinuity 324 (±10) km above the CMB (2567 km deep) with changes of V_S (δV_S) and V_P (δV_P) of 2.1% and −0.1%, respectively, from which we calculated a V_Φ change (δV_Φ) of −1.4%. A density contrast of 1.1% is assumed on the basis of associating the V_S discontinuity with the conversion from Pv to pPv; the data are insensitive to this choice, and there is no constraint on the density structure. The estimated uncertainties are ±0.2% for δV_S and ±0.15% for δV_P on the basis of the suites of models found to give adequate fits to the data (fig. S6B). Any sharp V_P discontinuity near 324 km above the CMB must have a δV_P <±0.15%, given the low noise levels of the stacks, which is a marked constraint on the structure. The V_S structure has some additional complexity at greater depths, especially near 191 km above the CMB. At this depth, the V_P structure does not have a discontinuity in our average models, but a δV_P of ~0.4% at variable depths can fit arrivals seen in smaller latitudinal bins (fig. S5, C and D). V_Φ decreases within the lowermost 40 km of the mantle, whereas V_P increases here.

For the 5° to 10° bin, the primary discontinuity is slightly deeper [299 km above the CMB (2592 km deep)], with δV_S = 1.1%, δV_P = −0.05%, and δV_Φ = −0.74%. The estimated δV_S depends on the choice of the direct S wave or ScS as a reference phase (we selected ScS) and could be 50% larger than the value given above (fig. S3), which leads to corresponding increases of the estimated δV_Φ and δV_Φ. A much stronger δV_P at this depth was proposed by Kito *et al.* (*18*).

The strongest feature in the P wave stacks for the 5° to 10° subregion is near 189 km above the CMB, where a change in velocity gradient with depth (dV_P/dz) and a 0.2 to 0.4% δV_P discontinuity are well resolved. The data sampling this subregion are at large distances where the sensitivity to this structure is high. The main effect comes from the change in dV_P/dz, as even smoothly varying models like Preliminary Reference Earth Model and International Association for Seismology and Physics of Earth's Interior model 1991 (IASP91) produce a general feature like that in the stacks (Fig. 3), but to match the stack amplitude, an additional small velocity jump is needed. This jump is sharp and must occur across <20 km in depth to fit data stacks at higher frequencies (fig. S6A). This feature was previously detected in (*17*). As noted above, V_S increased near this depth in the 10° to 15° subregion, but there is no clear structure in the 5° to 10° region. This structure is within the interval where pPv is likely to be present (Fig. 2). A pronounced V_Φ reduction occurs 51 km above the CMB with a small increase in V_P.

The velocity discontinuities near 300 km above the CMB in these models are sharp for simplicity; the V_S discontinuities have been shown to occur across no more than 30 km in depth (*20*). For both subregions, spreading the corresponding V_P discontinuity across 30 km in depth decreases the apparent reflectivity by no more than a factor of 2, and our uncertainty estimates allow for this. Whereas tight bounds are placed on the structure at this depth, the near-CMB structure is less well resolved.

To ensure that the waveform stacks are not contaminated by out-of plane scattering, we apply a Kirchhoff diffraction migration method (*13*), which seeks coherent scattering from a large 3D grid of possible isotropic point scatterers. This provides a more general parameterization of the medium than our localized 1D data stack-

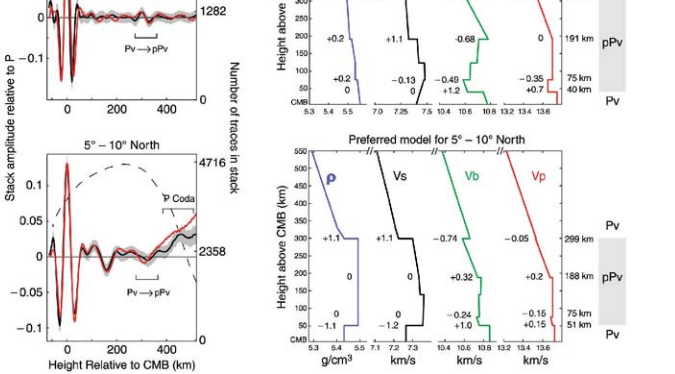


Fig. 2. **(A)** P wave double-array stacking results for the two 5° latitudinal bins indicated in Fig. 1. The solid black line shows the data-stack amplitude relative to the direct P wave amplitude. The gray shaded region represents the 95% confidence interval of the stack from bootstrap resampling, and the dashed line indicates the number of seismograms contributing to the data stack at a particular depth (scales at right). Above 400 km, the stack for 5° to 10°N is contaminated by P wave coda. The solid red line denotes stacks of synthetic seismograms generated from our preferred models. The synthetics were sampled to match the source-receiver geometry and had the same processing as the data. **(B)** Elasticity models obtained by modeling the data. Units are as follows: ρ, grams per cubic centimeter; V_Φ, kilometers per second; V_S, kilometers per second; and V_P, kilometers per second. Percent changes in these parameters are shown at first-order discontinuities at the indicated depths. The depth range likely to contain pPv is indicated at right.

ing (*19*). One profile through the data image volume (movie S1) is shown in Fig. 3, along with profiles through synthetic images. The migration images (Fig. 3) show regions of positive (red) and negative *P* wave arrival amplitudes (blue), with *PcP* forming a strong red image along the CMB with adjacent blue sidelobes. Suppressing the *PcP* energy in the data by zeroing amplitudes within 3 s of the *PcP* arrival before the migration allows weak shallower features to show up better. With the ray paths being limited to a single corridor, there are artifacts due to limited source-receiver geometry and narrow frequency bandwidth; however, these are well accounted for in the migration images of synthetic seismograms, with a comparison of the data and IASP91 synthetics highlighting features not explained by a smooth velocity model. Synthetics for our preferred 1D velocity models from the double-array stacking analysis clearly match the data well, particularly the structure ~300 km above the CMB (shown in blue in Fig. 3). The migrations do not show any evidence of contamination from out-of-plane scatterers, ensuring that the 1D models are reasonable for the ~300-km lateral Fresnel zones averaged by the data in the 0.25 to 0.5 Hz passband.

Globally, the lowermost 800 km of the mantle is observed to have large-scale heterogeneities in seismic velocities, and V_s and V_b are anticorrelated overall (*21*–*23*), but this is at least partly the result of large low-shear velocity provinces (LLSVPs) under the Pacific and Africa having negative V_s anomaly with weaker negative or no V_p anomaly, which is not the behavior expected for pPv. The LLSVP observations are commonly attributed to a chemical anomaly (*24*–*26*); if pPv is present in these regions (*27*), thermal and chemical effects must overwhelm the velocity effects of the phase change. Some regions in the lowermost 300 km of the mantle—notably beneath the circum-Pacific band of subduction zones—have positive V_s anomalies but limited resolution of V_p anomalies (*8*, *28*). The Cocos region is unique in that it permits V_s and V_p to be well enough resolved to directly compute V_b behavior reliably.

Our determination of anticorrelation of V_s and V_b under the Cocos, arising from the presence of a strong V_s increase with a small V_p decrease, is consistent with the expectations for the Pv-to-pPv transition, supporting this interpretation of the structure. The overall velocities in our models are faster than those in global models, which indicates lower-than-average temperatures favorable to the presence of pPv, but the phase change itself appears to cause most of the V_s anomaly. For this region, then, a laboratory calibration of temperatures at the phase transition pressure indicated by the discontinuity provides a reasonable estimate of absolute mantle *T* of ~2500 K at the discontinuity, and properties of pPv can be considered to explain the overall seismic structure above the CMB (*15*, *20*). This includes the reductions of V_s right above the CMB in our models and the accompanying increases in V_p, which can be explained if a steep thermal gradient reverts pPv to Pv right above the CMB (*15*, *27*, *29*).

Some experiments and theories suggest that increasing the Fe content in Pv lowers the pressure at which Pv transforms to pPv (*6*, *30*, *31*); however, other high *P*-*T* experiments with pyrolitic-type compositions (~10% Fe content) show very limited differences in transition pressure relative to end member $MgSiO_3$ when different pressure standards are used (*32*). The presence of Al is predicted to increase the Pv-to-pPv transition pressure and to broaden the pressure range of the two-phase zone up to 13 GPa (~150 km) (*33*, *34*), which would dramatically weaken any seismic energy reflected from the phase change. Including either Fe or Al in pPv tends to reduce the changes in seismic properties across the Pv-to-pPv transition, with the effect of Fe being about three times stronger than that of Al (*34*, *35*). Although estimates of the precise properties of pPv differ between methods, for pyrolitic-type silicates it is consistently found that the jump in V_s is strong and positive, any change in V_p is small, and the predicted change in V_b is strong and negative, consistent with our observations under the Cocos plate. Whereas complex compositional variations (such as former oceanic crust, if it has sunk to the base of the mantle) might still be needed to explain some of the secondary features in the Cocos region, the structure near 2600 km of depth in this region is well explained by the Pv-to-pPv phase change.

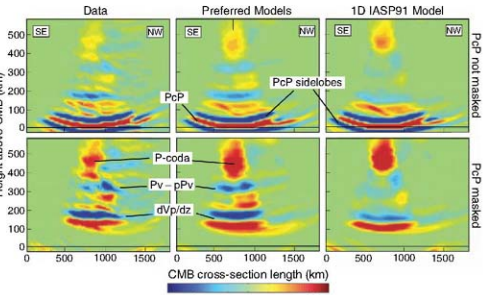


Fig. 3. Vertical cross sections through the volume imaged by scattering migration for the *P* wave data set with *PcP* as the reference phase. Horizontal axes denote lateral distance along the line in Fig. 1 going from southeast (SE) to northwest (NW). The CMB is shown by the black line. The top row is for the full waveforms, and the bottom is for waveforms with *PcP* masked out with a 6-s-wide symmetric taper applied around the predicted arrival time for *PcP* for each seismogram. The left column shows the migration stack of observed data, and the other columns show migration images of synthetic seismograms generated from a merge of synthetics for our preferred models (Fig. 2) for the two subregions (middle) and the 1D model (IASP91) (right). The synthetic merging was based on whether the *PcP* reflection point at the CMB was north or south of 10°N (indicated with a black tick mark in the upper middle panel). All data stacks have been weighted proportional to the number of seismograms contributing to each grid point, so lateral transitions to zero values (green) represent very small data sampling, not the absence of structure.

References and Notes

1. M. Murakami, K. Hirose, K. Kawamura, N. Sata, Y. Ohishi, *Science* **304**, 855 (2004); published online 8 April 2004, 10.1126/science.1095932.
2. A. R. Oganov, S. Ono, *Nature* **430**, 445 (2004).
3. T. Itaka, K. Hirose, K. Kawamura, M. Murakami, *Nature* **430**, 442 (2004).
4. T. Tsuchiya, J. Tsuchiya, K. Umemoto, R. M. Wentzcovitch, *Earth Planet. Sci. Lett.* **224**, 241 (2004).
5. R. M. Wentzcovitch, T. Tsuchiya, J. Tsuchiya, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 543 (2006).
6. S. Stackhouse, J. P. Brodholt, in *Post-perovskite: The Last Mantle Phase Change*, K. Hirose, J. P. Brodholt, T. Lay, D. Yuen, Eds. (American Geophysical Union, Washington, DC, 2007), pp. 99–113.
7. J. Wookey, S. Stackhouse, J.-M. Kendall, J. Brodholt, G. D. Price, *Nature* **438**, 1004 (2005).
8. T. Lay, E. J. Garnero, in *The State of the Planet: Frontiers and Challenges in Geophysics*, R. S. J. Sparks, C. J. Hawkesworth, Eds. (American Geophysical Union, Washington, DC, 2004), pp. 25–41.
9. S. P. Grand, *Philos. Trans. R. Soc. London Ser. A* **360**, 475 (2002).
10. D. Sun, T.-R. A. Song, D. Helmberger, *Geophys. Res. Lett.* **33**, L12507 (2006).
11. T. Lay, E. J. Garnero, S. A. Russell, *Geophys. Res. Lett.* **31**, L15612 (2004).
12. L. Boschi, T. W. Becker, B. Steinberger, *Geochem. Geophys. Geosyst.* **8**, Q10006 (2007).
13. A. R. Hutko, T. Lay, E. J. Garnero, J. Revenaugh, *Nature* **441**, 333 (2006).
14. C. Thomas, E. J. Garnero, T. Lay, *J. Geophys. Res.* **109**, B08307 (2004).
15. R. D. van der Hilst *et al.*, *Science* **315**, 1813 (2007).
16. X. M. Ding, D. V. Helmberger, *Phys. Earth Planet. Inter.* **101**, 245 (1997).

17. C. Reasoner, J. Revenaugh, *J. Geophys. Res.* **104**, 955 (1999).
18. T. Kito, S. Rost, C. Thomas, E. J. Garnero, *Geophys. J. Int.* **169**, 631 (2007).
19. Additional details of the processing steps are available as supporting material on *Science* Online.
20. T. Lay, *Geophys. Res. Lett.* **35**, L03304 (2008).
21. G. Masters, G. Laske, H. Bolton, A. M. Dziewonski, in *Earth's Deep Interior: Mineral Physics and Tomography From the Atomic to the Global Scale*, S.-I. Karato, A. M. Forte, R. C. Liebermann, G. Masters, L. Stixrude, Eds. (American Geophysical Union, Washington, DC, 2000), pp. 63–87.
22. S. Ni, D. V. Helmberger, J. Tromp, *Geophys. J. Int.* **161**, 283 (2005).
23. L. Wen, P. Silver, D. James, R. Kuehnel, *Earth Planet. Sci. Lett.* **189**, 141 (2001).
24. J. Trampert, F. Deschamps, J. Resovsky, D. Yuen, *Science* **306**, 853 (2004).
25. A. K. McNamara, S. Zhong, *Nature* **437**, 1136 (2005).
26. E. J. Garnero, T. Lay, A. McNamara, in *The Origin of Melting Anomalies: Plates, Plumes and Planetary Processes*, G. R. Foulger, D. M. Jurdy, Eds. (Geological Society of America, Boulder, CO, 2007), pp. 79–101.
27. T. Lay, J. Hernlund, E. J. Garnero, M. S. Thorne, *Science* **314**, 1272 (2006).
28. R. L. Saltzer, R. D. van der Hilst, H. Kárason, *Geophys. Res. Lett.* **28**, 1335 (2001).
29. J. W. Hernlund, C. Thomas, P. J. Tackley, *Nature* **434**, 882 (2005).
30. W. L. Mao *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 9751 (2005).
31. R. Caracas, R. E. Cohen, *Geophys. Res. Lett.* **32**, L16310 (2005).
32. K. Hirose, R. Sinmyo, N. Sata, Y. Ohishi, *Geophys. Res. Lett.* **33**, L01310 (2006).
33. S. Akber-Knutson, G. Steinle-Neumann, P. D. Asimow, *Geophys. Res. Lett.* **32**, L14303 (2005).
34. T. Tsuchiya, J. Tsuchiya, *Geophys. Res. Lett.* **33**, L12504 (2006).
35. R. Caracas, R. E. Cohen, in *Post-perovskite: The Last Mantle Phase Transition*, K. Hirose, J. P. Brodholt, T. Lay, D. Yuen, Eds. (American Geophysical Union, Washington, DC, 2007), pp. 115–128.
36. We thank Q. Williams and J. Hernlund for helpful discussion and comments. This work was supported in part by NSF under grants EAR-0453804 and EAR-0453944. Seismic data were obtained from the Incorporated Research Institutions for Seismology, Berkeley, and Caltech/U.S. Geological Survey TRInet data centers.

Supporting Online Material

www.sciencemag.org/cgi/content/full/320/5879/1070/DC1
Materials and Methods
Figs. S1 to S6
Movie S1

29 January 2008; accepted 8 April 2008
10.1126/science.1155822

Differential Rescue of Light- and Food-Entrainable Circadian Rhythms

Patrick M. Fuller, Jun Lu, Clifford B. Saper*

When food is plentiful, circadian rhythms of animals are powerfully entrained by the light-dark cycle. However, if animals have access to food only during their normal sleep cycle, they will shift most of their circadian rhythms to match the food availability. We studied the basis for entrainment of circadian rhythms by food and light in mice with targeted disruption of the clock gene *Bmal1*, which lack circadian rhythmicity. Injection of a viral vector containing the *Bmal1* gene into the suprachiasmatic nuclei of the hypothalamus restored light-entrainable, but not food-entrainable, circadian rhythms. In contrast, restoration of the *Bmal1* gene only in the dorsomedial hypothalamic nucleus restored the ability of animals to entrain to food but not to light. These results demonstrate that the dorsomedial hypothalamus contains a *Bmal1*-based oscillator that can drive food entrainment of circadian rhythms.

The circadian timing system (CTS) exerts endogenous temporal control over a wide range of physiological and neurobehavioral variables, conferring the adaptive advantage of predictive homeostatic regulation (*1*). When food is freely available, light signals from the retina entrain circadian rhythms to the day-night cycle (*2*). However, when food is available only during the normal sleep period [restricted feeding (RF)], many of these rhythms are reset so that the active phase corresponds to the period of food availability (*3*, *4*). In light entrainment, retinal ganglion cells that contain the photopigment melanopsin provide signals to neurons in the suprachiasmatic nuclei (SCN) of the hypothalamus that generate circadian rhythms through a series of molecular transcriptional, translational, and posttranslational feedback loops (*5*). The SCN in turn synchronizes circadian rhythms in peripheral tissue clocks through synaptic and humoral outputs (*5*, *6*). During RF, the SCN remains on the light cycle and SCN lesions do not prevent food entrainment, which suggests that another clock may supersede the SCN (*3*, *4*, *7*). Two recent studies have suggested that at least one food-entrainable clock may be located in the dorsomedial nucleus of the hypothalamus (DMH), but the importance of this clock for food entrainment has been disputed (*3*, *8*–*10*).

The core components of the molecular clock include the activating transcription factors BMAL1 and CLOCK and the negative regulatory feedback elements encoded by the *Per* and *Cry* genes (*11*, *12*). *Bmal1* is the only circadian clock gene without a known functional paralog and hence the only one for which a single gene deletion causes a complete loss of behavioral and molecular rhythmicity (*13*). Because its gene product BMAL1 is a transcription factor that likely influences many downstream genes, *Bmal1*$^{-/-}$ mice also exhibit other physiological defects unrelated to the circadian defect (*14*), including progressive arthropathy, decreased locomotor activity levels and body mass, and a shortened life span (*15*–*18*).

In this study, we used *Bmal1*$^{-/-}$ mice, which harbor a null allele at the *Bmal1* locus (*19*). The circadian patterns of locomotor activity (LMA) and body temperature (T_b) were monitored by telemetry (Fig. 1, A to C, and fig. S2). As previously reported, these animals showed no circadian rhythms in a 12-hour-light/12-hour-dark (LD) cycle or constant darkness (DD) when given ad libitum (AL) access to food (Fig. 1B). We also attempted to entrain *Bmal1*$^{-/-}$ mice to a 4-hour window of RF during the normal sleep period for mice, under both LD [ZT4-8 (from 4 to 8 hours after light onset)] and DD [CT4-8 (from 4 to 8 hours after presumptive light onset)] conditions. In LD and DD conditions, wild-type (WT) and heterozygous littermates showed an elevation of T_b and LMA ~2 to 3 hours before food became available (Fig. 2, A and D, and fig S2). By contrast, *Bmal1*$^{-/-}$ mice did not show a comparable elevation in T_b or increase in LMA before the window of RF in DD; T_b and LMA were, however, markedly elevated in the *Bmal1*$^{-/-}$ mice after food presentation (Fig. 2, B and D, Fig. 3C, and fig. S2). In addition to the prepprandial elevation in T_b under RF, WT and heterozygous littermates demonstrated a clear circadian T_b rhythm (Fig. 2A), whereas *Bmal1*$^{-/-}$ mice demonstrated a persisting ultradian T_b pattern throughout the remainder of the day (Fig. 2B). In DD conditions, *Bmal1*$^{-/-}$ mice occasionally showed periods of torpor (T_b below 31°C), which were distributed randomly across the circadian day. Consequently, the *Bmal1*$^{-/-}$ mice not only failed to show elevation of T_b or LMA in anticipation of the RF but often slept or were in torpor through the window of RF, requiring us to arouse them by gentle handling after presentation of the food to avoid their starvation and death during RF.

After 14 days in this RF regimen, mice were killed to examine clock gene expression in the brain and were compared to mice that had been fed AL. As previously reported (*8*), WT animals with AL food showed peak expression of *Per1* and *Per2* mRNA at ZT5-6, and *Bmal1* at ZT18-19 in the SCN (Fig. 3, D and E), but little or no expression at other hypothalamic sites. By contrast, WT animals under RF also showed no change in this expression pattern in the SCN (Fig. 3, D and E) but did show induction of *Per1* and *Per 2* at ~ZT3-9 (preceding, during, and after the RF window) in the DMH with peak expression levels at ZT7-8 (Fig. 3F). We also saw induction of *Bmal1* mRNA in the DMH, with peak expression at ZT18-21 (Fig. 3G), consistent with neurons in the DMH showing induction of rhythmic

Department of Neurology, Division of Sleep Medicine, and Program in Neuroscience, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA 02215, USA.

*To whom correspondence should be addressed. E-mail: csaper@bidmc.harvard.edu

expression of the entire suite of clock genes during RF. As previously reported for *Per1* and *Per2*, *Bmal1* gene expression was restricted to the compact region of the DMH. Finally, as expected in *Bmal1*$^{-/-}$ mice, no *Bmal1* mRNA and very low expression levels of *Per1* and *Per2* were detected in the SCN and DMH under any feeding or lighting condition (Fig. 1E, Fig. 2, E to G, and fig. S4).

These results suggest that a BMAL1-based circadian clock may be induced in the DMH after starvation and refeeding to drive entrainment of circadian rhythms to the time of food availability.

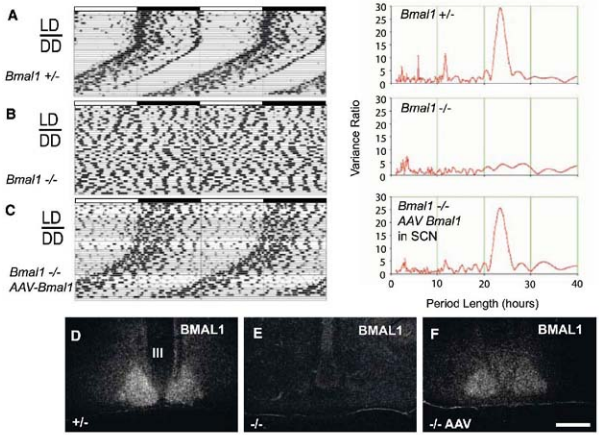


Fig. 1. Bmal1 expression in SCN rescues circadian rhythms and light entrainment in *Bmal1*$^{-/-}$ mice. Representative body temperature (T_b) records (double-plotted actigrams) and period analyses for (**A**) *Bmal1*$^{+/-}$, (**B**) *Bmal1*$^{-/-}$, and (**C**) *Bmal1*$^{-/-}$ mice with AAV-BMAL1 injections into the SCN under both light-dark (LD) and constant darkness (DD) conditions with AL access to food. The power spectrum for period (right panel) shows a dominant frequency at 23.7 hours for both Bmal1 heterozygous littermates and mice with AAV-BMAL1 rescue in the SCN, and the lack of a ~24-hour harmonic in the *Bmal1*$^{-/-}$ mice. (**D** to **F**) *Bmal1* mRNA expression (by in situ hybridization) in the SCN in (D) *Bmal1*$^{+/-}$ mice, (E) *Bmal1*$^{-/-}$ mice, and (F) after AAV-BMAL1 rescue by injection into the SCN (all ZT18). Scale bar, 100 μm.

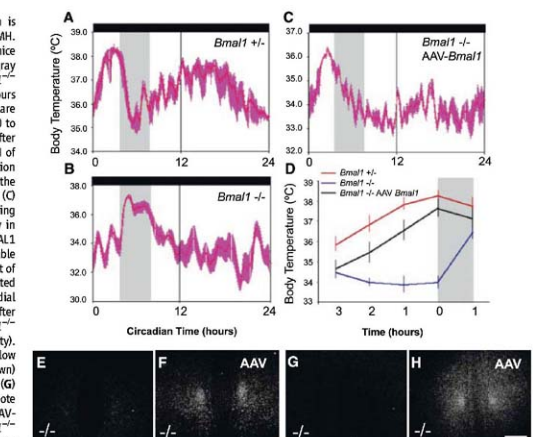


Fig. 2. Food entrainment of the T_b rhythm is rescued by AAV-BMAL1 injection into the DMH. Under DD conditions, *Bmal1*$^{+/-}$ and *Bmal1*$^{-/-}$ mice were placed under RF (food available CT4-8, gray vertical bar). (**A**) *Bmal1*$^{+/-}$ but not (**B**) *Bmal1*$^{-/-}$ mice demonstrated a clear rise in T_b ~2 to 3 hours before food availability under RF [(A) to (C) are waveforms showing mean ± SEM T_b on days 10 to 14 of RF for an individual mouse]. (**C**) After bilateral injection of AAV-BMAL1 into the DMH of *Bmal1*$^{-/-}$ mice, there was a preprandial elevation in T_b, but not a circadian rise in T_b during the presumptive dark cycle (CT12-24) [compare (C) with (A)]. As seen in the summary data plot, during the 3-hour window preceding food availability in RF (**D**), *Bmal1*$^{-/-}$ mice with DMH AAV-BMAL1 injections ($n = 4$, black trace) showed a comparable preprandial elevation in T_b (mean ± SEM) to that of *Bmal1*$^{-/+}$ mice (red trace), whereas noninjected *Bmal1*$^{-/-}$ mice (blue trace) showed no preprandial elevation in T_b; an increase in T_b occurred only after food presentation in the noninjected *Bmal1*$^{-/-}$ mice (gray bar in D is first hour of food availability). Under RF, *Bmal1*$^{-/-}$ mice demonstrated very low expression levels of *Per1* mRNA (**E**) (CT4 shown) and no expression of *Bmal1* mRNA at CT18 (**G**) (near-peak expression time in the heterozygote littermates) in the DMH. By contrast, after AAV-BMAL1 injection into the DMH (**F** and **H**), *Bmal1*$^{-/-}$ mice demonstrated robust *Per1* expression in the DMH at CT4 (F) and *Bmal1* expression at CT18 (H).

This hypothesis is consistent with previous studies showing that the DMH is a major recipient of direct and relayed input from the SCN and that it is important in relaying circadian signals for sleep-wake cycles, LMA, feeding, and corticosteroid rhythms to other brain systems (*20*). c-Fos expression in the DMH, but not in the SCN, is shifted to coincide with the activation of LMA and T_b during RF, and lesions of the DMH prevented entrainment of LMA, T_b, and sleep-wake cycles to RF [(*3*); also see (*9*, *10*)].

To test the role of the DMH-inducible clock in entrainment of circadian rhythms, we attempted to rescue both light and food entrainment of circadian rhythms by injecting adeno-associated viral vectors (AAV, serotype 8) containing the *Bmal1* gene (including both 5′ and 3′ promoter elements) (*19*) into the brains of *Bmal1*$^{-/-}$ mice. To test this construct, we first injected AAV-BMAL1 into the SCN. All mice with SCN injections ($n = 6$) (Fig. 1C), but none in which the injections missed the SCN ($n = 16$) or in which we injected a different AAV containing the gene for green fluorescent protein (GFP) into the SCN ($n = 3$), demonstrated entrainment of both LMA and T_b rhythms to a 12:12 LD cycle (Fig. 1C). When the mice were in DD conditions, this rhythm continued as a high-amplitude, free-running rhythm of 23.7 hours (Fig. 1C and fig. S1). Thus, focal bilateral injection of AAV-BMAL1 into the SCN of *Bmal1*$^{-/-}$ mice rescued the fundamental properties of the circadian oscillator, including light entrainment, free-running period, and rhythm amplitude. A previous study had shown that rescue of *Bmal1*$^{-/-}$ mice could be achieved by a transgene in which BMAL1 was placed under a constitutively active cytomegalovirus promoter (*13*). However, this gene construct, which was expressed continuously throughout the brain, produced a circadian period of ~22.7 hours, about 1 hour shorter than that of WT mice (*13*). By contrast, in our study, when we placed BMAL1 under its own promoter and restored this gene only to the SCN, the period of the circadian cycle was precisely the same as in WT animals (Fig. 1, A and C, right panel). Thus the expression of BMAL1 under its own promoter in the SCN alone is sufficient for recovery of light-entrained circadian rhythms. Our results also establish the SCN as sufficient for the generation of the circadian T_b rhythm, a point that has been in dispute (*21*).

On the other hand, similar to the mice with transgenic replacement of BMAL1 throughout the brain (*14*), locomotor activity levels in the animals with AAV-BMAL1 injections into the SCN remained significantly lower ($P < 0.001$) than those of WT and heterozygous littermate mice. Moreover, mice with AAV-BMAL1 injections into the SCN did not show improvement in any of the other physiologic deficits. Hence, these deficits are unlikely to be due to loss of circadian rhythmicity per se (*14*).

We next tested the animals with BMAL1 replacement in the SCN for their ability to entrain to a restricted temporal window of food availability. Previous studies had demonstrated that animals with lesions of the SCN could still entrain to food, suggesting that there was a food-entrainable oscillator elsewhere in the animals but not excluding participation of the SCN in intact animals (*4*, *7*, *22*). When animals with SCN injections of AAV-BMAL1 who had complete rescue of the light-entrained rhythms of

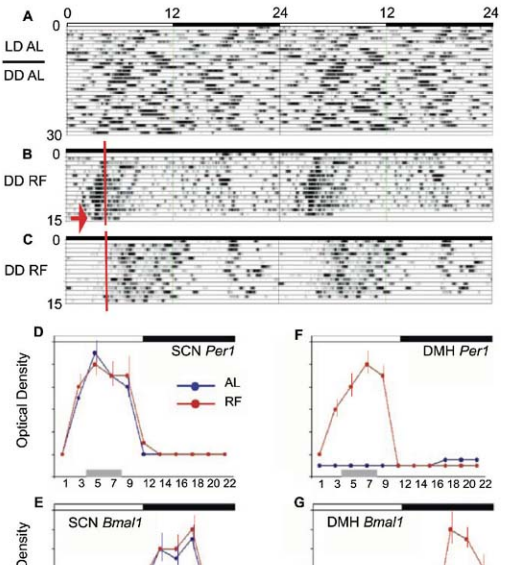


Fig. 3. AAV-Bmal1 injection into the DMH rescues food entrainment but not light entrainment in *Bmal1*$^{-/-}$ mice. (**A**) A double-plotted actigram of T_b from a *Bmal1*$^{-/-}$ mouse after bilateral injections of AAV-Bmal1 into the DMH. These mice demonstrate the lack of entrainment to the light-dark (LD) cycle and persisting ultradian rhythmicity in constant darkness (DD) under AL feeding conditions (Fig. 1B). These same mice, however, demonstrated anticipation and entrainment (**B**) to a RF cycle (red line, food presentation) in DD conditions (arrow, last day, no food given). By contrast, noninjected *Bmal1*$^{-/-}$ mice (**C**) did not demonstrate anticipation or entrainment to the RF cycle. (**D** to **G**) *Per1* and *Bmal1* gene expression (optical density, mean ± SEM; $n = 3$ mice per time point) in WT mice across the circadian day under both AL and RF (all data 12:12 LD) at 2-hour intervals, except for a 3-hour interval between ZT9-12. (D) *Per1* expression in the SCN showed a peak at ~ ZT6 under both AL and RF (gray bar, ZT4-8) conditions, demonstrating that the SCN remained phase-locked to the LD cycle during RF. (E) *Bmal1* expression in the SCN showed a peak at ~ZT18 under both AL and RF, further indicating that the SCN remained synchronized to the LD cycle during RF. (F) *Per1* expression in the DMH was undetectable at all ZT under AL; by contrast, *Per1* expression was sharply up-regulated by RF, with a peak at ~ZT7-8. (G) *Bmal1* expression in the DMH was also undetectable at all ZT (except very modest expression ~ZT 3-5) under AL, however, and similar to *Per1*, during RF *Bmal1* demonstrated up-regulation with a peak at ~ ZT18.

LMA and T_b were placed into a food-restriction paradigm under DD conditions, we found that they maintained the rhythm that had been entrained by light (high-amplitude, free-running period of ~23.7 hours) and never showed an increase in LMA or T_b in anticipation of the food presentation (fig. S3). Hence, although a BMAL1-based clock is necessary to support food entrainment, restoration of clock function in the SCN alone is not able to rescue this behavior.

To test the hypothesis that the BMAL1-based clock induced in the DMH during restricted feeding might drive circadian entrainment, we performed stereotaxic bilateral delivery of AAV-BMAL1 (the same construct and vector as used in the SCN) into the DMH of *Bmal1*$^{-/-}$ mice. Mice who sustained bilateral DMH injections of the AAV-BMAL1 did not demonstrate entrainment to a 12:12 LD cycle or free-running rhythms of T_b or LMA in DD (Fig. 3A). By contrast, under conditions of food restriction in DD, they exhibited a clear anticipatory increase in T_b and LMA before food presentation (Fig. 2C and Fig. 3B). Each individual mouse showed very little day-to-day variation in the timing of the increase in T_b and LMA under DD (i.e., the phase angle of entrainment was stable). Finally, the increase in T_b and LMA before the predicted period of food presentation persisted during a 24-hour fast at the end of restricted feeding (arrow in Fig. 3B), demonstrating the circadian nature of the response.

In both our study and the study by Mieda *et al.* (*8*), clock gene expression in the DMH was largely restricted to cells in the compact part of the nucleus, which consists of small, closely packed neurons that are highly reminiscent of the SCN itself. These neurons appear mainly to have local connections with the adjacent output zones of the DMH (*23*), suggesting that the timing signal from the compact part of the DMH may impinge upon the same output neurons in the remainder of the DMH as are used to control light-entrained rhythms directed by the SCN. This relationship may explain how the DMH clock is able to override the SCN clock input during conditions of food entrainment in an intact animal. It is unlikely that feedback from the DMH alters activity in the SCN in any major way, because the SCN remains phase-locked to the LD cycle for many weeks during food entrainment (as long as the animals are not also hypocaloric). These observations also raise the interesting possibility that the DMH may form the neuroanatomic basis of the so-called methamphetamine-sensitive circadian oscillator (MASCO), which also operates independent of the SCN and does not entrain to light [for a review, see (*24*)].

Our data indicate that there is an inducible clock in the DMH that can override the SCN and drive circadian rhythms when the animal is faced with limited food availability. Thus, under restricted feeding conditions, the DMH clock can assume an executive role in the temporal regulation of behavioral state. For a small mammal, finding food on a daily basis is a critical mission. Even a few days of starvation, a common threat in natural environments, may result in death. Hence, it is adaptive for animals to have a secondary "master clock" that can allow the animal to switch its behavioral patterns rapidly after a period of starvation to maximize the opportunity of finding food sources at the same time on following days.

In an intact animal, peripheral oscillators in many tissues in the body, including the stomach and liver, as well as elsewhere in the brain, may contribute to food entrainment of circadian rhythms (*25*, *26*). Consequently, it has been difficult to dissect this system by using lesions of individual components of the pathway (*3*, *9*, *10*). However, by starting with a genetically arrhythmic mouse and using focal genetic rescue in the brain, we have identified the SCN molecular clock as sufficient for light but not food entrainment of T_b and LMA rhythms in mice, and the DMH as sufficient for food but not light entrainment of circadian rhythms of T_b and LMA. These results demonstrate the power of viral-based gene replacement in the central nervous system to dissect complex neural functions.

References and Notes

1. M. H. Hastings, A. B. Reddy, E. S. Maywood, *Nat. Rev. Neurosci.* **4**, 649 (2003).
2. C. B. Saper, J. Lu, T. C. Chou, J. Gooley, *Trends Neurosci.* **28**, 152 (2005).
3. J. J. Gooley, A. Schomer, C. B. Saper, *Nat. Neurosci.* **9**, 398 (2006).
4. F. K. Stephan, J. M. Swann, C. L. Sisk, *Behav. Neural Biol.* **25**, 346 (1979).
5. P. L. Lowrey, J. S. Takahashi, *Annu. Rev. Genomics Hum. Genet.* **5**, 407 (2004).
6. A. Kramer *et al.*, *Science* **294**, 2511 (2001).
7. F. K. Stephan, *Physiol. Behav.* **46**, 489 (1989).
8. M. Mieda *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 12150 (2006).
9. G. J. Landry, M. M. Simons, I. C. Webb, R. E. Mistlberger, *Am. J. Physiol.* **290**, 6 (2006).
10. J. J. Gooley, C. B. Saper, *J. Biol. Rhythms* **22**, 479 (2008).
11. N. Gekakis *et al.*, *Science* **280**, 1564 (1998).
12. B. Zheng *et al.*, *Cell* **105**, 683 (2001).
13. M. K. Bunger *et al.*, *Cell* **103**, 1009 (2000).
14. E. L. McDearmon *et al.*, *Science* **314**, 1304 (2006).
15. M. K. Bunger *et al.*, *Genesis* **41**, 122 (2005).
16. A. Laposky *et al.*, *Sleep* **28**, 395 (2005).
17. S. Shimba *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 12071 (2005).
18. R. D. Rudic *et al.*, *PLoS Biol.* **2**, e377 (2004).
19. Materials and methods are available as supporting material on *Science* Online.
20. T. C. Chou *et al.*, *J. Neuroscience* **23**, 1069 (2003).
21. C. A. Fuller *et al.*, *Am. J. Physiol.* **241**, 5 (1981).
22. D. T. Krieger, H. Hauser, L. C. Krey, *Science* **197**, 398 (1977).
23. R. H. Thompson, N. S. Canteras, L. W. Swanson, *Brain Res. Brain Res. Rev.* **27**, 89 (1998).
24. T. Hiroshige, K. Honma, S. Honma, *Brain Res. Bull.* **27**, 441 (1991).
25. K. A. Stokkan *et al.*, *Science* **291**, 490 (2001).
26. F. Damiola *et al.*, *Genes Dev.* **14**, 2950 (2000).

27. We thank Q. Ha and M. Ha for technical work. Support was provided by U.S. Public Health Service grants HL60292, NS33987, NS051609, NS057119, and HL07901-08.

Supporting Online Material

www.sciencemag.org/cgi/content/full/320/5879/1074/DC1
Materials and Methods
Figs. S1 to S4
References

26 November 2007; accepted 14 April 2008
10.1126/science.1153277

Endogenous siRNAs Derived from Transposons and mRNAs in *Drosophila* Somatic Cells

Megha Ghildiyal,[1]* Hervé Seitz,[1]* Michael D. Horwich,[1] Chengjian Li,[1] Tingting Du,[1] Soohyun Lee,[2] Jia Xu,[3] Ellen L.W. Kittler,[4] Maria L. Zapp,[4] Zhiping Weng,[2] Phillip D. Zamore[1]†

Small interfering RNAs (siRNAs) direct RNA interference (RNAi) in eukaryotes. In flies, somatic cells produce siRNAs from exogenous double-stranded RNA (dsRNA) as a defense against viral infection. We identified endogenous siRNAs (endo-siRNAs), 21 nucleotides in length, that correspond to transposons and heterochromatic sequences in the somatic cells of *Drosophila melanogaster*. We also detected endo-siRNAs complementary to messenger RNAs (mRNAs); these siRNAs disproportionately mapped to the complementary regions of overlapping mRNAs predicted to form double-stranded RNA in vivo. Normal accumulation of somatic endo-siRNAs requires the siRNA-generating ribonuclease Dicer-2 and the RNAi effector protein Argonaute2 (Ago2). We propose that endo-siRNAs generated by the fly RNAi pathway silence selfish genetic elements in the soma, much as Piwi-interacting RNAs do in the germ line.

Three RNA-silencing pathways have been identified in flies and mammals: RNA interference (RNAi), guided by small interfering RNAs (siRNAs) derived from exogenous double-stranded RNA (dsRNA); the microRNA (miRNA) pathway, in which endogenous small RNAs repress partially complementary mRNAs; and the Piwi-interacting RNA (piRNA) pathway, whose small RNAs repress transposons in the germ line (*1*–*3*) and can activate transcription in heterochromatin (*4*).

Endogenous siRNAs (endo-siRNAs) silence retrotransposons in plants (*5*, *6*), and siRNAs corresponding to the L1 retrotransposon have

[1]Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, MA 01605, USA. [2]Program in Bioinformatics, Boston University, Boston, MA 02215, USA. [3]Department of Biomedical Engineering, Boston University, Boston, MA 02215, USA. [4]Program in Molecular Medicine and Center for AIDS Research, University of Massachusetts Medical School, Worcester, MA 01605, USA. [5]Program in Bioinformatics and Integrative Biology, University of Massachusetts Medical School, Worcester, MA 01605, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: phillip.zamore@umassmed.edu

been detected in cultured mammalian cells (*7*). Genetic and molecular evidence suggests that in addition to suppressing viral infection, the RNAi pathway silences selfish genetic elements in the fly soma: Mutations in the RNAi gene *rm62* (*8*) suppress mutations caused by retroelement insertion (*9*); depletion of the Argonaute proteins Ago1 or Ago2 increases transposon expression in cultured *Drosophila* Schneider 2 (S2) cells (*10*); small RNAs have been detected in *Drosophila* Kc cells for the *1360* transposon (*11*) and are produced during transgene silencing in flies (*12*); and siRNAs have been proposed to repress germline expression of *suffix*, a short interspersed nuclear element (SINE) (*13*).

The defining properties of *Drosophila* siRNAs are their production from long dsRNA by Dicer-2 (Dcr-2), which generates 5′-monophosphate termini; their loading into Argonaute2 (Ago2); and their Ago2-dependent, 3′-terminal, 2′-*O*-methylation by the methyltransferase Hen1 (*14*–*16*), unlike most miRNAs (*17*). In vivo (Fig. 1A, rightmost

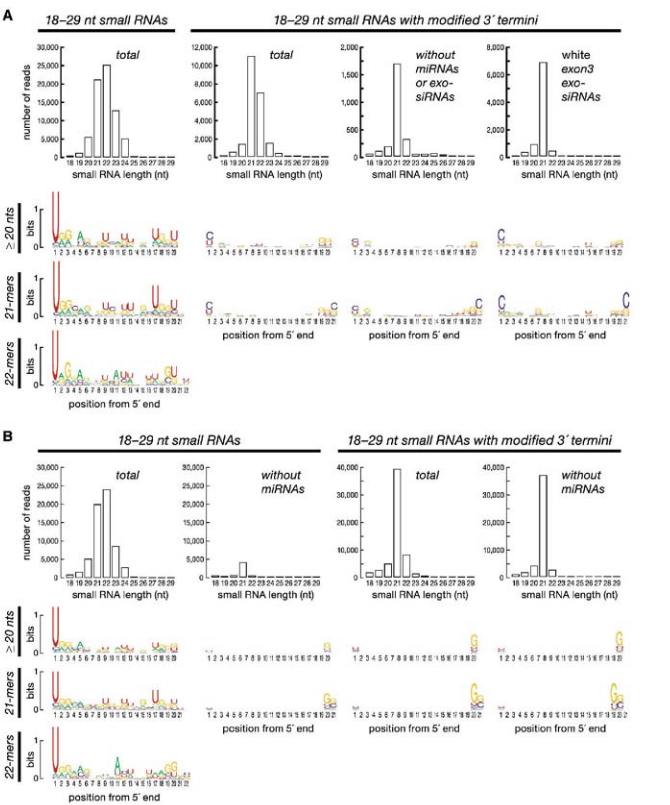


Fig. 1. High-throughput pyrosequencing revealed 3′-terminally modified 21-nt RNAs in the fly soma. **(A)** Length and sequence composition of the small RNA sequences from a library of total small RNA from the heads of flies expressing an inverted repeat (IR) silencing the *white* gene and for a parallel library enriched for RNAs modified at their 3′ ends. **(B)** Similar analysis for small RNA sequences from *Drosophila* S2 cells. For data labeled "without miRNAs," pre-miRNA–matching sequences were removed computationally.

panel) and in vitro (*18*), nearly all siRNAs produced by Dcr-2 from exogenous dsRNA are 21 nucleotides (nt) in length.

We characterized the somatic small RNA content of S2 cells (*19*) and of heads expressing an RNA hairpin silencing the *white* gene by RNAi (*20*). To identify endo-siRNA candidates, we analyzed two types of RNA libraries. For total 18- to 29-nt RNA libraries, 89% (S2 cells) and 96% (heads) mapped to annotated miRNA loci. In contrast, libraries enriched for small RNAs bearing a 3′-terminal, 2′-*O*-methyl modification (*21*) were depleted of miRNAs: Only 19% (S2 cells) and 49% (heads) of reads and 2.4% (S2 cells; 58,681 reads; 12,036 sequences) and 12% (heads; 22,685 reads; 2929 sequences) of unique sequences mapped to miRNA loci.

Table 1. Endo-siRNAs preferentially map to overlapping, complementary mRNAs.

Sample	Enrichment	Enrichment after randomization		*Z* score	*P*
		Mean	SD		
Fly heads	10.9	1.0	0.38	26.1	7.9×10^{-151}
S2 cells	12.3	1.1	0.42	27.0	5.2×10^{-161}

Figure 1 shows the length distribution and sequence composition of the four libraries. The total RNA samples were predominantly miRNAs, a bias reflected in their modal length (22 nt) and pronounced tendency to begin with uracil. Exclusion of miRNAs revealed a class of small RNAs with a narrow length distribution and no tendency to begin with uracil. Except for an unusual cluster of X-chromosome small RNAs (fig. S1) and a miRNA-like sequence with an unusual putative precursor on chromosome 2 (fig. S2), few of these small RNAs are likely to correspond to novel miRNAs: None lie in the arms of hairpins predicted to be as thermodynamically stable as most pre-miRNAs (i.e., < -15 kcal/mol).

After excluding known miRNAs, 64% (heads) (Fig. 1A) and 78% (S2 cells) (Fig. 1B) of sequences in the libraries enriched for 3′-terminally modified small RNAs—that is, those likely to be Ago2-associated—were 21 nt long. For fly heads, 37% (8404 reads) derived from the *white* dsRNA hairpin. The abundance of these exo-siRNAs can be estimated by comparing them to the number of reads for individual miRNAs in the total small RNA library, where 1.6% (660 antisense and 491 sense reads) were 21-nt oligomers (21-mers) and matched the *white* sequences in the dsRNA-expressing transgene. The collective abundance of all *white* exo-siRNAs was less than the individual abundance of the 10 most abundant miRNAs in this sample; the median abundance of any one exo-siRNA species was two reads. The *white*–inverted repeat (IR) transgene phenocopies a nearly null mutation in *white*, yet the sequence of the most abundant exo-siRNA was read just 37 times.

In heads, the sequence composition of the 21-nt, 3′-terminally modified small RNAs closely resembled that of exo-siRNAs, which tended to begin and end with cytosine. In heads and S2 cells, the 21-mers lacked the sequence features of piRNAs, which either begin with uracil (Aub- and Piwi-bound) or contain an adenine at position 10 (Ago3-bound) and are 23 to 29 nt long (*1*, *2*). These data suggest that the 21-nt small RNAs are somatic endo-siRNAs.

In S2 cells, endo-siRNAs mapped largely to transposons (86%); in fly heads, they mapped about equally to transposons, intergenic and unannotated sequences, and mRNAs. The finding that 41% of endo-siRNAs mapped to mRNAs without mapping to transposons suggests that endo-siRNAs may regulate mRNA

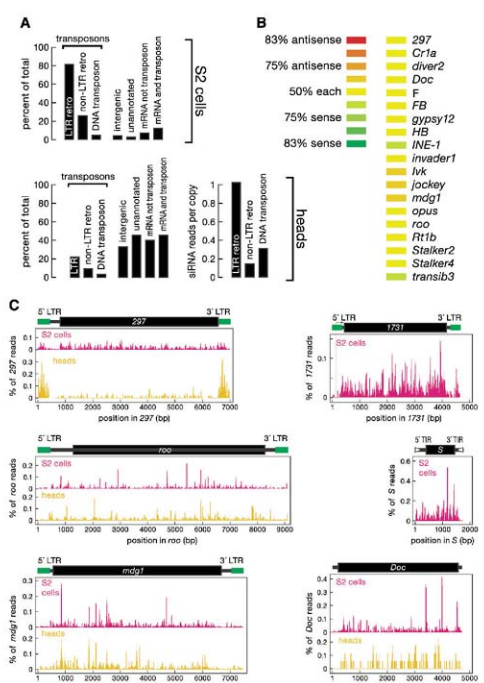


Fig. 2. Endo-siRNAs correspond to transposons. **(A)** Distribution of annotations for the genomic matches of endo-siRNA sequences. Bars total more than 100% because some siRNAs match both LTR and non-LTR retrotransposons or match both mRNA and transposons. **(B)** Transposon-derived siRNAs with more than 50 21-nt reads mapped about equally to sense and antisense orientations. **(C)** Alignment of endo-siRNA sequences to *Drosophila* transposons. The abundance of each sequence is shown as a percentage of all transposon-matching siRNA sequences. LTR, long terminal repeat; TIR, terminal inverted repeat. Here and in subsequent figures, data from high-throughput pyrosequencing and sequencing-by-synthesis were pooled for wild-type heads.

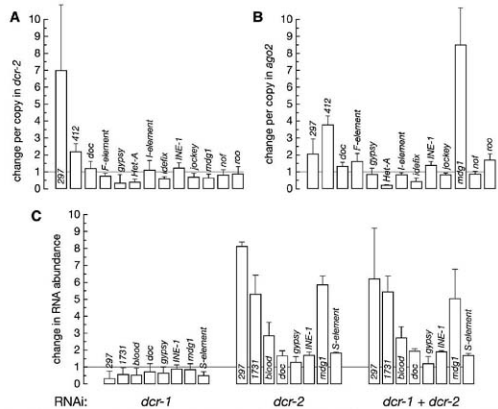


Fig. 3. Transposon silencing requires Dcr-2 and Ago2, but not Dcr-1. (**A** and **B**) The change in mRNA expression (mean ± SD, $N = 3$) for each transposon between $dcr\text{-}2^{L811fsX}$ (**A**) or $ago2^{414}$ (**B**) heterozygous and homozygous heads was measured by quantitative reverse transcription polymerase chain reaction. The data were corrected for differences in transposon copy number between the paired genotypes. (**C**) The change in transposon expression (mean ± SD, $N = 3$) in S2 cells was measured for the indicated RNAi depletion relative to a control dsRNA.

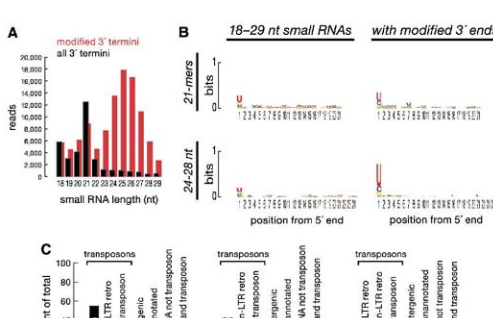


Fig. 4. The composition of somatic small RNAs is altered in the absence of Ago2. (**A** and **B**) Size distribution (**A**) and sequence composition (**B**) of sequences from a library of total 18- to 29-nt RNA from the heads of *ago2* null mutant flies or a library enriched for 3′-terminally modified RNAs. Reads matching pre-miRNA sequences were removed. (**C**) Distribution of annotations for the genomic matches of small RNA sequences from the two *ago2* libraries.

expression. Endo-siRNAs mapping to mRNAs were likelier by a factor of >10 than expected by chance ($5.22 \times 10^{-161} < P < 8 \times 10^{-151}$) to derive from genomic regions annotated to produce overlapping, complementary transcripts (Table 1 and table S1). These data suggest that such overlapping, complementary transcripts anneal in vivo to form dsRNA that is diced into endo-siRNAs. We note that among the mRNAs for which we detected complementary 21-mers was *ago2* itself.

Endo-siRNAs mapped to all three large chromosomes (figs. S3 to S5). siRNAs corresponding to the three transposon types in *Drosophila* were detected, but long terminal repeat (LTR) retrotransposons, the dominant class of selfish genetic elements in flies, were overrepresented even after accounting for their abundance in the genome (Fig. 2A and table S2). Unlike piRNAs, which are disproportionately antisense to transposons, but like siRNAs derived from exogenous dsRNA, about equal numbers of sense and antisense transposon-matching endo-siRNAs were detected (Fig. 2B and fig. S6) (*1*–*3*, *22*). Like piRNAs, endo-siRNAs map to large genomic clusters (table S3). Of 172 endo-siRNA clusters in S2 cells, four coincided with previously identified piRNA clusters (cluster 1, at 42A of chromosome 2R; clusters 7 and 10 in unassembled genomic sequence; and cluster 15 in the chromosome 3L heterochromatin). In heads, we detected 17 clusters; five corresponded to clusters found in S2 cells, but only one was shared with the germline piRNAs: the *flamenco* locus, consistent with recent genetic evidence that a Piwi-independent but *flamenco*-dependent pathway represses the *Idefix* and *ZAM* transposons in the soma (*23*). That both endo-siRNAs and piRNAs can arise from the same region suggests either that a single transcript can be a substrate for both piRNA and siRNA production or that distinct classes of transcripts arise from a single locus. The abundance and distribution of endo-siRNAs across the sequences of individual transposon species reflected the natural history of when the elements entered the fly genome, but not their mechanism of transposition (Fig. 2C) (*24*).

Statistically significant reductions in siRNA abundance were observed in $dcr\text{-}2^{L811fsX}$ null mutant heads relative to heads from heterozygous siblings for 38 transposons (fig. S7 and table S4). Normalized for sequencing depth, sequencing results from homozygous *dcr-2* mutant heads yielded fewer 21-mers overall (by a factor of 3.1) and fewer 21-mers corresponding to transposons (by a factor of 6.3) than did their heterozygous siblings ($P < 2.2 \times 10^{-16}$; χ^2 test). In contrast, overall miRNA abundance—normalized to sequencing depth—was essentially unchanged between *dcr-2* heterozygotes and homozygotes (fig. S7 and table S5). These data suggest that endo-siRNAs are produced by Dcr-2, but we do not yet know why some endo-siRNAs persist in $dcr\text{-}2^{L811fsX}$ mutants.

Transposon expression in the soma reflects both the silencing of transposons—potentially by either or both posttranscriptional and transcriptional mechanisms—and the tissue specificity of transposon promoters. *Drosophila* somatic cells may contain siRNAs targeting transposons that would not be highly expressed even in the absence of those siRNAs, because the promoters of those transposons are not active in some or all somatic tissues or because they are repressed by additional mechanisms. We analyzed the expression of a panel of transposons in heads from *ago2* and *dcr-2* mutants and in S2 cells depleted of Dcr-1, Dcr-2, or Ago2 by RNAi (Fig. 3 and fig. S8). We found that the steady-state abundance of RNA from the LTR retrotransposons *297* and *412* increased in heads from $dcr\text{-}2^{L811fsX}$ null mutants (Fig. 3A). Similarly, the steady-state abundance of RNA from the LTR retrotransposons *297*, *412*, *mdg1*, and *roo*, the non-LTR retrotransposon *F-element*, and the SINE-like element *INE-1* increased in $ago2^{414}$ mutant heads (Fig. 3B).

In S2 cells, RNA expression from the LTR retrotransposons *297*, *1731*, *mdg1*, *blood*, and *gypsy* and from the DNA transposon *S-element* all increased significantly ($0.00001 < P < 0.002$) when Dcr-2 was depleted or when both Dcr-2 and Dcr-1 were depleted, but not when Dcr-1 alone was depleted (Fig. 3C). Similarly, *ago2(RNAi)* in S2 cells desilenced transposons, including nine LTR and non-LTR retrotransposons and the DNA transposon *S-element* (fig. S8).

Is Ago2 required for the production or accumulation of endo-siRNAs? We sequenced 18- to 29-nt small RNAs from $ago2^{414}$ homozygous fly heads and from the same small RNA sample treated to enrich for 3′-terminally modified RNAs. After computationally removing miRNAs, the sequences from the untreated library contained a prominent 21-nt peak (Fig. 4A) that predominantly began with uracil (Fig. 4B), much like miRNAs and unlike siRNAs in wild-type heads, which often began with cytosine (Fig. 1A). Perhaps in the absence of Ago2, only a subpopulation of endo-siRNAs that can bind Ago1 accumulates. The small RNAs from the $ago2^{414}$ library enriched for 3′-terminally modified sequences were predominantly 24 to 27 nt long and often began with uracil—a length distribution and sequence bias characteristic of piRNAs, which, like siRNAs, are 2′-*O*-methylated at their 3′ ends. Both the 21-nt small RNAs and the piRNA-like RNAs in the *ago2* mutant heads mapped to transposons, unannotated heterochromatic and unassembled sequences, but the piRNA-like sequences mapped to mRNAs far less frequently than did either the 21-mers or wild-type endo-siRNAs (Fig. 4C). How these piRNA-like small RNAs are generated and whether they contribute to transposon silencing in the fly soma remain unknown.

Note added in proof: The loci described here in figs. S1 and S2 correspond to endo-siRNA–generating hairpins recently identified in (*25*–*27*).

References and Notes

1. L. S. Gunawardane *et al.*, *Science* **315**, 1587 (2007); published online 21 February 2007 (10.1126/science.1140494).
2. J. Brennecke *et al.*, *Cell* **128**, 1089 (2007).
3. V. V. Vagin *et al.*, *Science* **313**, 320 (2006); published online 28 June 2006 (10.1126/science.1129333).
4. H. Yin, H. Lin, *Nature* **450**, 304 (2007).
5. A. Hamilton, O. Voinnet, L. Chappell, D. Baulcombe, *EMBO J.* **21**, 4671 (2002).
6. R. Sunkar, T. Girke, J. K. Zhu, *Nucleic Acids Res.* **33**, 4443 (2005).
7. N. Yang, H. H. J. Kazazian, *Nat. Struct. Mol. Biol.* **13**, 763 (2006).
8. A. Ishizuka, M. C. Siomi, H. Siomi, *Genes Dev.* **16**, 2497 (2002).
9. A. K. Csink, R. Linsk, J. A. Birchler, *Genetics* **138**, 153 (1994).
10. J. Rehwinkel *et al.*, *Mol. Cell. Biol.* **26**, 2965 (2006).
11. K. A. Haynes, A. A. Caudy, L. Collins, S. C. Elgin, *Curr. Biol.* **16**, 2222 (2006).
12. M. Pal-Bhadra, U. Bhadra, J. A. Birchler, *Mol. Cell* **9**, 315 (2002).
13. N. A. Tchurikov, O. V. Kretova, *PLoS ONE* **2**, e476 (2007).
14. M. D. Horwich *et al.*, *Curr. Biol.* **17**, 1265 (2007).
15. A. Pelisson, E. Sarot, G. Payen-Groschene, A. Bucheton, *J. Virol.* **81**, 1951 (2007).
16. K. Saito *et al.*, *Genes Dev.* **21**, 1603 (2007).
17. K. Okamura, A. Ishizuka, H. Siomi, M. C. Siomi, *Genes Dev.* **18**, 1655 (2004).
18. A. Nykanen, B. Haley, P. D. Zamore, *Cell* **107**, 309 (2001).
19. *Drosophila* RNAi Screening Center at Harvard Medical School (http://flyrnai.org/cgi-bin/RNAi_FAQ_lines.pl).
20. Y. S. Lee, R. W. Carthew, *Methods* **30**, 322 (2003).
21. H. Seitz, M. Ghildiyal, P. D. Zamore, *Curr. Biol.* **18**, 147 (2008).
22. P. D. Zamore, T. Tuschl, P. A. Sharp, D. P. Bartel, *Cell* **101**, 25 (2000).
23. S. Drosot, N. Buchon, C. Meignin, M. Coiffet, C. Vaury, *PLoS ONE* **3**, e1526 (2008).
24. See supporting material on *Science* Online.
25. B. Czech *et al.*, *Nature* 10.1038/nature07007 (2008).
26. Y. Kawamura *et al.*, *Nature* 10.1038/nature06938 (2008).
27. K. Okamura *et al.*, *Nature* 10.1038/nature07015 (2008).

28. We thank A. Boucher and S. Ma for technical assistance; G. Farley for encouragement, support, and technical assistance; and Roche Applied Science for high-throughput sequencing. P.D.Z. is a W. M. Keck Foundation Young Scholar in Medical Research. Supported by NIH grants GM62862 and GM65236 (P.D.Z.), GM080625 (J.X. and Z.W.), and HG003367 (S.L.); EMBO long-term (ALTF 910-2004) and Human Frontier Science Program (LT00575/2005-L) fellowships (H.S.); and a National Research Service Award predoctoral MD/PhD fellowship from the National Institute on Aging (F30AG030283) (M.D.H.). NCBI Gene Expression Omnibus accession numbers for sequence and abundance data are GSE9389 and GSE11019, respectively.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1157396/DC1
Materials and Methods
Figs. S1 to S8
Tables S1 to S7
References

5 March 2008; accepted 31 March 2008
Published online 10 April 2008;
10.1126/science.1157396
Include this information when citing this paper.

Resource Partitioning and Sympatric Differentiation Among Closely Related Bacterioplankton

Dana E. Hunt,[1]* Lawrence A. David,[2]* Dirk Gevers,[1,3,4] Sarah P. Preheim,[1] Eric J. Alm,[1,5]† Martin F. Polz[1]†

Identifying ecologically differentiated populations within complex microbial communities remains challenging, yet is critical for interpreting the evolution and ecology of microbes in the wild. Here we describe spatial and temporal resource partitioning among *Vibrionaceae* strains coexisting in coastal bacterioplankton. A quantitative model (AdaptML) establishes the evolutionary history of ecological differentiation, thus revealing populations specific for seasons and life-styles (combinations of free-living, particle, or zooplankton associations). These ecological population boundaries frequently occur at deep phylogenetic levels (consistent with named species); however, recent and perhaps ongoing adaptive radiation is evident in *Vibrio splendidus*, which comprises numerous ecologically distinct populations at different levels of phylogenetic differentiation. Thus, environmental specialization may be an important correlate or even trigger of speciation among sympatric microbes.

Microbes dominate biomass and control biogeochemical cycling in the ocean, but we know little about the mechanisms and dynamics of their functional differentiation in the environment. Culture-independent analysis typically reveals vast microbial diversity, and although some taxa and gene families are differentially distributed among environments (*1*, *2*), it is not clear to what extent coexisting genotypic diversity can be divided into functionally cohesive populations (*1*, *3*). First, we lack broad surveys of nonpathogenic free-living bacteria that establish robust associations of individual strains with spatiotemporal conditions (*4*, *5*); second, it remains controversial what level of genetic diversification reflects ecological differentiation. Phylogenetic clusters have been proposed to correspond to ecological populations that arise by neutral diversification after niche-specific selective sweeps (*6*). Clusters are indeed observed among closely related isolates (e.g., when examined by multilocus sequence analysis) (*7*) and in culture-independent analyses of coastal bacterioplankton (*8*). Yet recent theoretical studies suggest

that clusters can result from neutral evolution alone (*9*), and evidence for clusters as ecologically distinct populations remains sparse, having been most conclusively demonstrated for cyanobacteria along ocean-scale gradients (*10*) and in a depth profile of a microbial mat (*11*). Further, horizontal gene transfer (HGT) may erode the ecological cohesion of clusters if adaptive genes are transferred (*12*), and recombination can homogenize genes between ecologically distinct populations (*13*). Thus, exploring the relationship between phylogenetic and ecological differentiation is a critical step toward understanding the evolutionary mechanisms of bacterial speciation (*9*).

In this study, we investigated ecological differentiation by spatial and temporal resource partitioning in coastal waters among coexisting bacteria of the family *Vibrionaceae*, which are ubiquitous, metabolically versatile heterotrophs (*14*). The coastal ocean is well suited to test population-level effects of microhabitat preferences, because tidal mixing and oceanic circulation ensure a high probability of migration, reducing biogeographic effects on population structure. In the plankton, heterotrophs may adopt alternate ecological strategies: exploiting either the generally lower concentration but more evenly distributed dissolved nutrients or attaching to and degrading small suspended organic particles, originating from algal exopolysaccharides and detritus (*3*). Bacterial microhabitat preferences may develop because resources are distributed on the same scale as the dispersal range of individuals, due to turbulent mixing and active motility (*15*). Of potential microhabitats, particles represent abundant but relatively short-lived resources, as labile components are rapidly utilized (on time scales of hours to days) (*16*, *17*), implying that particle colonization is a dynamic process. Moreover, particulate matter may change composition with macroecological conditions (such as seasonal algal blooms). Zooplankton provide additional, more stable microhabitats; vibrios attach to and metabolize chitinous zooplankton exoskeletons (*18*, *19*) but may also live in the gut or occupy niches specific to pathogens. The extent to which microenvironmental preferences contribute to resource partitioning in this complex ecological landscape remains an important question in microbial ecology (*20*).

We aimed to conservatively identify ecologically coherent groups by examining distribution patterns of *Vibrionaceae* genotypes among free-living and associated (with suspended particles and zooplankton) compartments of the plankton-

[1]Department of Civil and Environmental Engineering, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139, USA. [2]Computational and Systems Biology Initiative, MIT, Cambridge, MA 02139, USA. [3]Laboratory of Microbiology, Ghent University, Gent 9000, Belgium. [4]Bioinformatics and Evolutionary Genomics Group, Ghent University, Gent 9000, Belgium. [5]Department of Biological Engineering, MIT, Cambridge, MA 02139, USA, and Broad Institute of MIT and Harvard University, Cambridge, MA 02139, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: ejalm@mit.edu (E.J.A.); mpolz@mit.edu (M.F.P.)

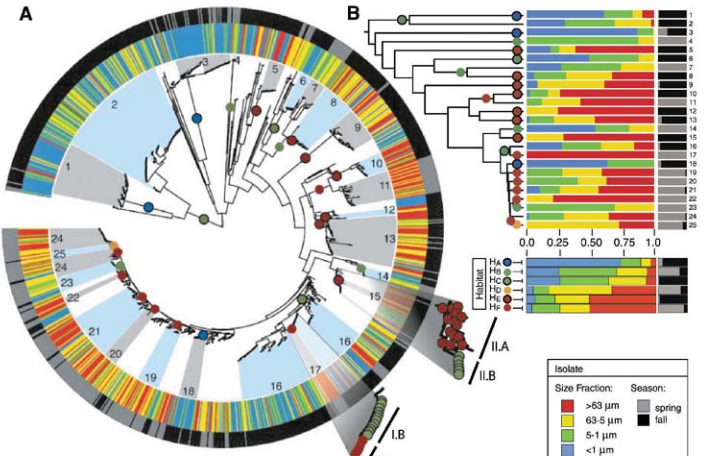


Fig. 1. Season and size fraction distributions and habitat predictions mapped onto *Vibrionaceae* isolate phylogeny inferred by maximum likelihood analysis of partial *hsp60* gene sequences. Projected habitats are identified by colored circles at the parent nodes. **(A)** Phylogenetic tree of all strains, with outer and inner rings indicating seasons and size fractions of strain origin, respectively. Ecological populations predicted by the model are indicated by alternating blue and gray shading of clusters if they pass an empirical confidence threshold of 99.99% (see SOM for details). Bootstrap confidence levels are shown in fig. S10. **(B)** Ultrametric tree summarizing habitat-associated populations identified by the model and the distribution of each population among seasons and size fractions. The habitat legend matches the colored circles in (A) and (B) with the habitat distribution over seasons and size fractions inferred by the model. Distributions are normalized by the total number of counts in each environmental category to reduce the effects of uneven sampling. The insets at the lower right of (A) show two nested clusters (I.A and I.B and II.A and II.B) for which recent ecological differentiation is inferred, including habitat predictions at each node. The closest named species to numbered groups are as follows: G1, *V. calviensis*; G2, *Enterovibrio norvegicus*; G3, *V. ordalii*; G4, *V. rumoiensis*; G5, *V. alginolyticus*; G6, *V. aestuarianus*; G7, *V. fischerivlogei*; G8, *V. fischeri*; G9, *V. superstes*; G10, *V. penaeicida*; G11 to G25, *V. splendidus*.

ic environment under different macroecological conditions (spring and fall) (fig. S1 and table S1). Because the level of genetic differentiation at which ecological preferences develop is not known, we focused on a range of relationships [0 to 10% small subunit ribosomal RNA (rRNA) divergence] among co-occurring vibrios (*21*). Particle-associated and free-living cells were separated into four consecutive size fractions by sequential filtration (four replicate water samples, each subsampled with at least four replicate filters per size fraction); each fraction contained organisms and dead organic material of different origins [detailed in the supporting online material (SOM)]. For simplicity, we refer to these fractions as enriched in zooplankton (≥63 μm), in large (5 to 63 μm) and small (1 to 5 μm) particles, and in free-living cells (0.22 to 1 μm) (fig. S1B). The 1- to 5-μm size fraction was somewhat ambiguous, probably containing small particles as well as large or dividing cells; however, it provided a firm buffer between obviously particle-associated (>5 μm) and free-living (<1 μm) cells. *Vibrionaceae* strains were isolated by plating filters on selective media, previously shown by quantitative polymerase chain reaction to yield good correspondence between genotypes recovered in culture and those present in environmental samples (*21*). Roughly 1000 isolates were characterized by partial sequencing of a protein-coding gene (*hsp60*). To obtain added resolution, between one and three additional gene fragments

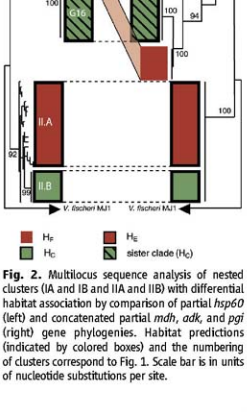


Fig. 2. Multilocus sequence analysis of nested clusters (IA and IB and IIA and IIB) with differential habitat association by comparison of partial *hsp60* (left) and concatenated partial *mdh*, *adk*, and *pgi* (right) gene phylogenies. Habitat predictions (indicated by colored boxes) and the numbering of clusters correspond to Fig. 1. Scale bar is in units of nucleotide substitutions per site.

(*mdh*, *adk*, and *pgi*) were sequenced for over half of the isolates (SOM), including *V. splendidus* strains, the most abundant group (*21*).

Our rationale for testing environmental associations grows out of the following considerations. First, as in most ecological sampling, the true habitats or niches are unknown and can only be observed as projections onto the sampling dimensions ("projected habitats"). Thus, associations can be detected as distinct distributions of groups of strains if habitats/niches are differentially apportioned among samples. Second, the lack of an accepted microbial species concept implies that it is imprudent to use any measure of genetic relationships to define a priori the populations whose environmental association should be assessed. Therefore, we first tested the null hypothesis that there is no environmental association across the phylogeny of the strains. We then refined such estimates by developing a new model to simultaneously identify populations and their projected habitats. Finally, these model-based results were tested with nonparametric empirical statistics.

The initial null hypothesis of no association between phylogeny and ecology is strongly rejected (seasons: $P < 10^{-79}$; size fractions: $P < 10^{-49}$) by comparing the parsimony score of observed environments on the tree to that expected by chance (*22*) (SOM), confirming the visual impression of differential patterns of clustering among seasons and size fractions (Fig. 1A). This result is robust toward uncertainty in the phylogeny, which should diminish but not strengthen associations, and is confirmed by introducing additional uncertainty in the phylogeny (fig. S2). The observed overall association with season and size fraction therefore suggests that water-column vibrios partition resources, but neither provides insights into the phylogenetic bounds of populations or the composition of their habitats.

We therefore developed an evolutionary model (AdaptML) to identify populations as groups of related strains sharing a common projected habitat, which reflects their relative abundance in the measured environmental categories (size fractions and seasons) (SOM). In practice, the model inputs are the phylogeny, season, and size fraction of the strains. It then maps changes in environmental preference onto the tree by predicting projected habitats for each extant and ancestral strain in the phylogeny. Although similar in spirit to existing parsimony, likelihood, and Bayesian methods, which map ancestral states onto trees (*23*), the model accounts for the complexities and uncertainties of environmental sampling. First, projected habitats can span multiple sampling dimensions to account for complex life cycles (such as time spent in multiple true habitats) and problems inherent in environmental sampling: Discrete samples rarely equate to true habitats, and true habitats are frequently misplaced among their typical sample categories (for example, zooplankton fragments may also be found in smaller size fractions). Second, projected habitats can span multiple phylogenetic clusters to allow for the possibility that clusters may arise neutrally or that the relevant parameters differentiating them ecologically have not been measured.

Briefly, AdaptML builds a hidden Markov model for the evolution of habitat associations: Adjacent nodes on the phylogeny transition between habitats according to a probability function that is dependent on branch length and a transition rate, which is learned from the data (SOM) (fig. S3). Subsequently, we optimize the model parameters (the transition rate and the composition of each projected habitat) to maximize the likelihood of the observed data. Finally, we use a simple ad hoc rule for reducing noninformative parameters: We merge habitats that converge to similar distributions (simple correlation of distribution vectors >90%) during the model-fitting procedure (SOM). This reproducibly identified six nonredundant habitats for the observed data set (H_A to H_F in Fig. 1B and fig. S5). Moreover, the algorithm acts conservatively, as suggested by two tests. First, the model did not overfit the data when there was no ecological signal present: When the environments were shuffled, only a single generalist habitat (evenly distributed over all size fractions and seasons) was recovered. Second, when simulated habitats were used to generate environmental assignments, the model usually identified a number of habitats equal to or less than the true number present (fig. S6).

The analysis suggests that a single bacterial family coexisting in the water column resolves into a striking number of ecologically distinct populations with clearly identifiable preferences (habitats). The algorithm identified 25 populations, associated with one of the six habitats defined by distinct distributions of isolates over seasons and size fractions (Fig. 1 and fig. S7). Most clusters have a strong seasonal signal; interestingly, two pairs of highly similar habitats are observed in both seasons (Fig. 1B). The first of the habitat pairs corresponds to populations occurring both free-living and on particles but lacking zooplankton-associated isolates (H_B and H_C); the second indicates a preference for zooplankton and large particles (H_E and H_F) (Fig. 1B). The remaining two habitats were season-specific. Habitat H_A combines all primarily free-living populations in the fall, whereas habitat H_D identifies a second particle- and zooplankton-associated group in spring, but unlike H_E and H_F it has a higher proportion of large particles and maps onto a single small group (G25) (Fig. 1). However, we cannot place high confidence in the absence of the free-living habitat in the spring, because relatively few strains were recovered from that fraction. Moreover, the distribution of individual populations among seasons and size fractions varies considerably, with remarkably narrow preferences for some populations whereas others are more broadly distributed. For example, *V. ordalii* (G3) is almost exclusively free-living in both seasons, whereas *V. alginolyticus* (G5) has a significant representation in both zooplankton and free-living size fractions but

occurs exclusively in the fall (Fig. 1, A and B). The sequences of three additional genes for *V. alginolyticus* isolates were identical, arguing against misidentification due to recombination or additional population substructuring. Similarly, there was good agreement when two different gene phylogenies (*hsp60* and *mdh*) were used to identify habitats for *V. splendidus* (fig. S8), although fewer habitats were identified using the *mdh* tree, most likely because it is less well-resolved. Overall, across all vibrios sampled, association with the zooplankton-enriched and free-living fractions dominated, and although several populations contain particle-associated isolates, only a few appear to be specifically particle-adapted. Because vibrios are generally regarded as particle and zooplankton specialists (*14*), this observed partitioning offers new insight into their ecology.

Thus, in spite of the highly variable conditions of the water column, populations appear to finely partition resources, especially because our habitat estimates are conservative, as clusters occupying the same habitat may be differentiated along additional (unobserved) resource axes. For example, different zooplankton-associated groups may be host- or body region–specific, and the strong seasonal signal of most clusters may be due to a variety of factors; however, temperature is a likely candidate because it has so far arisen as the strongest correlate of microbial population changes both over a seasonal cycle (*24*) and along ocean-scale gradients (*10*). Finally, populations, which appear unassociated in our study, may be true generalists with respect to the resource space sampled or may be adapted to environments not sampled in this study, such as animal intestines or sediments (*14*). Despite these uncertainties, the observed strong partitioning among associated and free-living clusters may have important implications for population biology in the bacterioplankton. As recently suggested (*9*), for attached bacteria, the effective population size (N_e) may be considerably smaller than the census size because colonization serves as a population bottleneck, whereas in free-living clusters, N_e may be closer to the census size. Although computing the true magnitude of N_e in microbial populations remains controversial (*25*), it is an important parameter that determines the relative strength of selection and drift. Thus, attached and free-living populations may evolve under different constraints (*9*).

The phylogenetic structure of populations also provides insights into the history of habitat switches. Deeply branching populations may have remained associated with habitats over long evolutionary time, and shallow branches may have diversified more recently (Fig. 1, A and B). These stable habitat-associated clusters roughly correlate to named species within the *Vibrionaceae*. For example, *V. ordalii* (G3) and *Enterovibrio norvegicus* (G2) both represent clusters without close relatives containing >50 isolates, which are overwhelmingly predicted to follow primarily free-living (H_A) and free-living/particle-associated life-styles (H_C), respectively (Fig. 1A). On the other hand, some very closely related clusters are associated with different habitats; *V. splendidus*, which is composed of strains that are ~99% identical in rDNA gene sequence (*21*), differentiates into 15 microdiverse habitat-associated clusters, of which one is distributed roughly evenly among both seasons, and 9 and 5 predominantly occur in spring and fall, respectively. Thus, *V. splendidus* appears to have ecologically diversified, possibly by invading new niches or partitioning resources at increasingly fine scales.

Recent or perhaps ongoing radiation by sympatric resource partitioning is most strongly suggested for two nested clusters within *V. splendidus*, where groups of strains differing by as little as a single nucleotide in *hsp60* display distinct ecological preferences (Fig. 1A, insets, and table S2). These strains were isolated from multiple independent samples and thus do not represent clonal expansion, suggesting that this may reflect a true habitat switch; nonetheless, homologous recombination could also move alleles between distantly related, ecologically distinct clusters, creating spurious phylogenetic relationships, which can be detected by comparison with other genes. Multilocus sequence analysis shows that for nested cluster I, a close relationship was artificially created because *hsp60* gene phylogeny is discordant with three other genes (Fig. 2). However, this still represents a habitat switch, just at a slightly larger sequence distance, as I.A is nested within the much larger G16 cluster in both the *hsp60* and the *mdh-pgi-adk* phylogenies. For the second nested cluster, the three additional genes confirm partial separation of the subclusters II.A and II.B by a single base pair difference in one of the genes, whereas the other genes consist of identical alleles. This reinforces the idea that subcluster II.A is not incorrectly grouped because of recombination, despite its distinct ecological affiliation (Fig. 2). In combination, these data support the idea that there is ecological differentiation among recently diverged genotypes and show that such changes might be recognized in protein-coding genes as soon as they accumulate (neutral) sequence changes.

How might adaptation to a new habitat relate to speciation, the generation of distinct clusters of closely related bacteria? Mathematical modeling has recently shown that the dynamics of speciation depend on the ratio of homologous recombination to mutation rates (r/m) (*9*). When this ratio per allele exceeds ~1, populations transition from essentially clonal to sexual, with the major consequence that selection is probably required for the formation of clusters (*9*). Our preliminary multilocus sequence analysis on a set of strains with similar taxonomic composition suggests that their r/m is well above that threshold. Thus, our observations of habitat separation for highly similar but clearly distinct genotypes suggest that ecological selection may have triggered phylogenetic differentiation. A plausible mechanism is that differential distribution among habitats (possibly caused by few adaptive loci) is sufficient to depress gene flow between associated genotypes (*9*, *26*). Consequently, mutations will no longer be homogenized but instead accumulate within specialized populations, even for ecologically neutral genes. Over time, genetic isolation may increase because homologous recombination rates decrease log-linearly with sequence distance (*27*). We detected associations with different habitats among sister clades over a wide range of phylogenetic distances, possibly representing populations at various stages of differentiation (Fig. 1A). Although we cannot determine whether clusters represent transiently adapted populations or nascent species, our observations of differential distributions of genotypes suggest that there exists a small-scale adaptive landscape in the water column allowing the initiation of (sympatric) speciation within this community.

Although it has recently been suggested that microbial lineages remain specific to macroenvironments over long evolutionary times (*28*), this study demonstrates switches in ecological associations within a bacterial family coexisting in the coastal ocean. In the *V. splendidus* clade, speciation could be ongoing, but the divergence between most other ecologically defined groups appears large. This is consistent with our previous suggestion that rRNA gene clusters, which are roughly congruent with the deeply divergent protein-coding gene clusters detected here, represent ecological populations (*8*). However, the example of *V. splendidus* highlights the fact that using marker genes to assess community-wide diversity may not capture some ecological specialization. Moreover, different groups of organisms could evolve under different constraints, and the mechanisms suggested here apply to the invasion of new habitats and are thus different from (but compatible with) the widely discussed niche-specific selective sweeps (*29*). Why *V. splendidus* appears to have radiated recently into new habitats whereas other groups appear to be more constant is not known but may be related to its high heterogeneity in genome architecture (*21*). This could indicate a large (flexible) gene pool that, if shared by horizontal gene transfer, gives rise to large numbers of ecologically adaptive phenotypes. It will therefore be important to compare whole genomes within recently ecologically diverged clusters to identify specific changes leading to adaptive evolution.

References and Notes

1. S. J. Giovannoni, U. Stingl, *Nature* **437**, 343 (2005).
2. E. F. DeLong *et al.*, *Science* **311**, 496 (2006).
3. M. F. Polz, D. E. Hunt, S. P. Preheim, D. M. Weinreich, *Philos. Trans. R. Soc. London Ser. B* **361**, 2009 (2006).
4. A. Ramette, J. M. Tiedje, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 2761 (2007).
5. J. B. H. Martiny *et al.*, *Nat. Rev. Microbiol.* **4**, 102 (2006).
6. F. M. Cohan, E. B. Perry, *Curr. Biol.* **17**, R373 (2007).
7. D. Gevers *et al.*, *Nat. Rev. Microbiol.* **3**, 733 (2005).
8. S. G. Acinas *et al.*, *Nature* **430**, 551 (2004).
9. C. Fraser, W. P. Hanage, B. G. Spratt, *Science* **315**, 476 (2007).

10. Z. L. Johnson *et al.*, *Science* **311**, 1737 (2006).
11. M. J. Ferris, M. Kühl, A. Wieland, D. M. Ward, *Appl. Environ. Microbiol.* **69**, 2893 (2003).
12. W. F. Doolittle, R. T. Papke, *Genome Biol.* **7**, 116 (2006).
13. A. C. Retchless, J. G. Lawrence, *Science* **317**, 1093 (2007).
14. J. R. Thompson, M. F. Polz, in *The Biology of Vibrios*, F. L. Thompson, B. Austin, J. Swings, Eds. (American Society for Microbiology Press, Washington, DC, 2006), pp. 190–203.
15. T. Kjærbye, H. P. Grossart, H. Ploug, K. Tang, *Appl. Environ. Microbiol.* **68**, 3996 (2002).
16. L. M. Pomeroy *et al.*, *Bull. Mar. Sci.* **35**, 426 (1984).
17. C. Panagiotopoulos *et al.*, *Organ. Geochem.* **33**, 985 (2002).
18. J. F. Heidelberg, K. B. Heidelberg, R. R. Colwell, *Appl. Environ. Microbiol.* **68**, 5498 (2002).
19. D. E. Hunt, D. Gevers, N. M. Vahora, M. F. Polz, *Appl. Environ. Microbiol.* **74**, 44 (2008).
20. J. Pernthaler, R. Amann, *Microbiol. Mol. Biol. Rev.* **69**, 440 (2005).
21. J. R. Thompson *et al.*, *Science* **307**, 1311 (2005).
22. W. P. Maddison, M. Slatkin, *Evolution* **45**, 1184 (1991).
23. T. Ronquist, *Trends Ecol. Evol.* **19**, 475 (2004).
24. J. R. Thompson *et al.*, *Appl. Environ. Microbiol.* **70**, 4103 (2004).
25. V. Daubin, N. A. Moran, *Science* **306**, 978 (2004).
26. F. M. Cohan, *Genetica* **116**, 359 (2002).
27. J. Majewski, *FEMS Microbiol. Let.* **199**, 161 (2001).
28. C. von Mering *et al.*, *Science* **315**, 1126 (2007).
29. A. Koeppel *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 2504 (2008).
30. We thank E. DeLong and D. Distel for critical reading of the manuscript. This work was supported by grants from the U.S. Department of Energy Genomes to Life program to M.F.P. and E.J.A. and by support from the NSF/National Institute of Environmental Health Sciences Woods Hole Center for Oceans and Human Health, the NSF Biological Oceanography Program, and the Moore Foundation to M.F.P. D.G. is indebted to the Fund for Scientific Research, Flanders (Belgium), for a postdoctoral fellowship and research funding. L.A.D. gratefully acknowledges support from a U.S. Department of Defense National Defense Science and Engineering Graduate fellowship. Sequences have been submitted to GenBank under accession nos. EU653713 to EU655560.

Supporting Online Material
www.sciencemag.org/cgi/content/full/320/5879/1081/DC1
Materials and Methods
Figs. S1 to S10
Tables S1 and S2
References

17 March 2008; accepted 18 April 2008
10.1126/science.1157890

A Polymorphism Within the *G6PC2* Gene Is Associated with Fasting Plasma Glucose Levels

Nabila Bouatia-Naji,[1]* Ghislain Rocheleau,[2]* Leentje Van Lommel,[3] Katleen Lemaire,[3] Frans Schuit,[3] Christine Cavalcanti-Proença,[1] Marion Marchand,[1] Anna-Liisa Hartikainen,[4] Ulla Sovio,[5] Franck De Graeve,[1] Johan Rung,[2] Martine Vaxillaire,[1] Jean Tichet,[6] Michel Marre,[7] Beverley Balkau,[8] Jacques Weill,[9] Paul Elliott,[5] Marjo-Riitta Jarvelin,[5,10] David Meyre,[1] Constantin Polychronakos,[2,11] Christian Dina,[1] Robert Sladek,[2] Philippe Froguel[1,12]†

Several studies have shown that healthy individuals with fasting plasma glucose (FPG) levels at the high end of the normal range have an increased risk of mortality. To identify genetic determinants that contribute to interindividual variation in FPG, we tested 392,935 single-nucleotide polymorphisms (SNPs) in 654 normoglycemic participants for association with FPG, and we replicated the most strongly associated SNP (rs560887, $P = 4 \times 10^{-7}$) in 9353 participants. SNP rs560887 maps to intron 3 of the *G6PC2* gene, which encodes glucose-6-phosphatase catalytic subunit–related protein (also known as IGRP), a protein selectively expressed in pancreatic islets. This SNP was associated with FPG (linear regression coefficient $\beta = -0.06$ millimoles per liter per A allele, combined $P = 4 \times 10^{-23}$) and with pancreatic β cell function (Homa-B model, combined $P = 3 \times 10^{-13}$) in three populations; however, it was not associated with type 2 diabetes risk. We speculate that *G6PC2* regulates FPG by modulating the set point for glucose-stimulated insulin secretion in pancreatic β cells.

Recent innovations in genotyping technology have led to the identification of genetic variants associated with increased risk of type 2 diabetes (T2D) (*1*–*4*). However, the genetic factors contributing to interindividual variation in blood glucose levels in the general population are largely unknown. In addition to affecting metabolic health, these variants may also affect coronary heart disease (CHD) risk, for which a linear correlation between FPG and CHD mortality has been described for both diabetic and nondiabetic individuals (*5*, *6*). Here, we describe the results of a genome-wide association (GWA) study in which we aimed to identify genetic variants involved in glucose homeostasis in the general population.

We analyzed FPG as a quantitative trait in 654 normoglycemic (NG) nonobese individuals by using Illumina Infinium HumanI (Illumina Incorporated, San Diego, CA) and Hap300 BeadArray (Illumina Incorporated, San Diego, CA) (table S1) (*1*, *7*). The strongest association signal was observed with SNP rs560887 [$P = 4 \times 10^{-7}$ adjusted for age, gender, and body mass index (BMI) under the additive model]. This SNP is part of a 17-kb linkage disequilibrium (LD) block on chromosome 2 that encompasses the gene and the 3′ flanking region of the islet-specific glucose-6-phosphatase–related protein [IGRP, also known by the gene name *G6PC2* (glucose-6-phosphatase catalytic unit 2)], a glycoprotein embedded in the endoplasmic reticulum (ER) membrane (*8*). Six SNPs within this LD block were genotyped in our GWA analysis. These included a nonsynonymous SNP (rs492594-L219V), which showed modest association with FPG [$P = 0.04$ for rs492594-L219V, which is in low LD with rs560887 at $r^2 = 0.24$ in the Centre d'Etude du Polymorphisme Humain (CEPH) families with Northern and Western European ancestry (CEU), according to HapMap data], and a SNP located in the 3′ flanking region of *G6PC2* (rs563694), which showed stronger association ($P = 2 \times 10^{-5}$, $r^2 = 0.76$ with rs560887) (table S2). We used linear regression analysis to adjust for the effect of rs560887 on all SNPs located within 100 kbp upstream and 100 kbp downstream that showed various levels of association with FPG ($P < 0.05$) and found no SNP that remained associated at significance level of 5% (table S2). Imputation of genotypes using the LD structure from HapMap data (CEU population) did not identify any SNP showing potentially stronger association with FPG (Fig. 1). On the basis of these findings, we tested the association of rs560887 with FPG by using a complementary subset (table S1) of the French DESIR (data from the Epidemiological Study on the Insulin Resistance Syndrome) population (DESIR NG stage 2, $N = 3419$) and replicated our original association with FPG (linear regression coefficient $\beta = -0.06$ mmol/l per A allele, $P = 3 \times 10^{-8}$) (Table 1).

IGRP has been proposed to modulate the glycolytic pathway and eventually glucose-stimulated insulin secretion by dephosphorylating glucose-6-phosphate generated by glucokinase, the β cell glucose sensor (*9*). *G6pc2* knockout mice display decreased FPG and normal insulin sensitivity (*10*). In the DESIR NG cohort, the rs560887 A allele [minor allele frequency (MAF) = 0.30] associated with decreased FPG is also associated with a decreased fraction of erythrocyte glyco-

[1]CNRS UMR 8090 Institute of Biology, Pasteur Institute of Lille and Lille 2 Droit et Santé University, 59019 Lille, France. [2]Department of Human Genetics, Faculty of Medicine, McGill University and Génome Québec Innovation Centre, Montreal H3A 1A4, Canada. [3]Gene Expression Unit, Department of Molecular Cell Biology, Katholieke Universiteit, B-3000 Leuven, Belgium. [4]Department of Obstetrics and Gynaecology, 90014 Oulu, Finland. [5]Department of Epidemiology and Public Health, Imperial College, London W2 1PG, UK. [6]Institut Inter Régional pour la Santé (IRSA), 37520 La Riche, France. [7]INSERM U695, Bichat Hospital, 75722 Paris, France. [8]INSERM U780-IFR69, Paris Sud University, 94807 Villejuif, France. [9]Pediatric Endocrine Unit, Jeanne de Flandre Hospital, 59037 Lille, France. [10]Department of Public Health and General Practice, University of Oulu, 90014 Oulu, Finland. [11]Department of Pediatrics, Faculty of Medicine, McGill University, Montreal H3H 1P3, Canada. [12]Genomic Medicine, Hammersmith Hospital, Imperial College London, London W12 0NN, UK.

*These authors equally contributed to this work.
†To whom correspondence should be addressed. E-mail: p.froguel@imperial.ac.uk

sylated hemoglobin (HbA1c) ($\beta = -0.04$, $P = 8 \times 10^{-5}$). This reflects a better long-term glucose control, with increased basal insulin secretion as assessed by the homeostasis model index of pancreatic β cell function Homa-B ($\beta = 5.46$, $P = 2 \times 10^{-5}$), a measure derived from physiological data of FPG and fasting insulin levels (*7*).

To validate these findings in different European populations, we genotyped young NG individuals from the Northern Finland Birth Cohort 86 (*7*) (NFBC86, $N = 5073$) and confirmed the association between rs560887 and decreased FPG ($\beta = -0.07$ mmol/l per A allele, $P = 6 \times 10^{-16}$) and increased Homa-B index ($\beta = 3.79$; $P = 6 \times 10^{-10}$). In contrast, no association was found between rs560887 and fasting insulin in either French ($P = 0.5$) or Finnish ($P = 0.1$) cohorts. Similarly, neither cohort showed strong association between rs560887 and BMI (DESIR $P = 0.1$, NFBC86 $P = 0.04$). Adjustment for BMI did not affect the association with FPG, suggesting that the observed effect on FPG is independent of obesity-induced insulin resistance. To test this hypothesis, we studied 861 severely obese NG French children (mean age 11 ± 3 years) (*7*). In this cohort, we still observed association between rs560887 and decreased FPG ($\beta = -0.07$ mmol/l per A allele, $P = 0.004$) and increased Homa-B index ($\beta = 9.13$, $P = 0.01$). As in the French DESIR and Finish cohorts, rs560887 did not associate with fasting insulin ($P = 0.30$) or BMI ($P = 0.60$) in obese children.

rs560887 is located in the third intron of *G6PC2*, 26 base pairs from its boundary with the fourth exon. Splice site modeling shows that the rs560887 A allele may have a lower splicing acceptor score than the G allele (www.cbcb.umd.edu/software/GeneSplicer) (*11*); however, further investigation is required to define the functional consequences of the observed association signal, which may be caused by rs560887 or an unidentified SNP in high LD with rs560887, and *G6PC2* splicing or gene expression.

Longitudinal analysis of the DESIR cohort showed that carriers of the rs560887 A allele had a decreased risk of developing mild hyperglycemia (defined as a FPG ≥ 6.1 mmol/l) over a 9-year follow-up period [hazard ratio = 0.83, 95% confidence interval (CI) 0.72 to 0.96, $P = 0.009$]. Surprisingly, rs560887, or any SNP within its 17-kb LD block, did not associate with T2D risk in our GWA study (*1*) or other published T2D case-control studies (www.broad.mit.edu/diabetes/ or www.wtcc.org.uk/). To confirm this, we analyzed 2972 additional T2D cases (table S1). With the 4073 DESIR NG individuals as controls, we observed no significant association with T2D risk ($P = 0.6$, odds ratio (OR) = 0.98 95% CI 0.91 to 1.05). We note that our case-control analysis achieved statistical power of at least 99.95% to detect a significant effect (MAF = 0.30, OR ≥ 1.20, $N = 7045$). Previous reports have shown that the SNP most strongly associated with T2D identified so far (rs7903146, located in *TCF7L2*) does not associate with FPG in European general populations (*12*, *13*), suggesting that the genetic determinants that regulate FPG in physiological states may be different from those that increase T2D risk. IGRP is homologous to two other ER proteins of the glucose-6-phosphatase family that catalyze the conversion of glucose-6-phosphate into glucose, a key enzymatic step

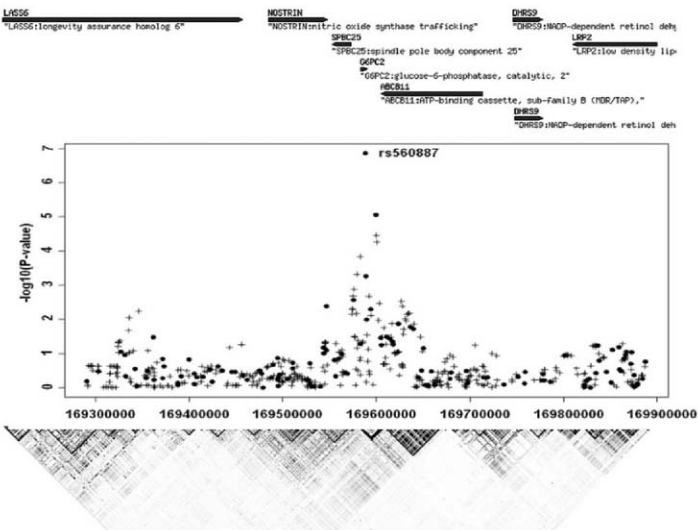


Fig. 1. Genomic context and association with FPG of the *G6PC2* gene variant rs560887 on chromosome 2q. **(Top)** The genomic context of rs560887 that maps to intron 3 of *G6PC2*. **(Middle)** The association magnitude [$-\log_{10}(P$ value)] with fasting plasma glucose of the SNPs from the Illumina 317 mapping array (Tag-SNPs) (dots) and the HapMap CEU data imputed SNPs (plus signs), not directly genotyped around (±100 kb) rs560887. **(Bottom)** The LD structure (defined by pairwise r^2) between SNPs in the CEU population from HapMap phase II using the Haploview software.

for both gluconeogenesis and glycogenolysis (*8*). In contrast to the hepatocyte-specific member (*G6pc1*) that is highly expressed in gluconeogenic tissues but not in pancreas (fig. S1, top), the expression of *G6pc2* is restricted to islets and pancreatic β cell line MIN6 (fig. S1, middle), as shown in humans (*14*). Because expression of *G6pc2* mRNA is restricted to β cells in the mouse, as indicated in previous studies (*14*) and as confirmed in our own experiments (fig. S1), rs560887 may play an important role in the early pathogenesis of this disease without contributing to overall T2D risk. For example, the rs560887 common allele (G) might alter the glucostat set point and up-regulate FPG, but the development of overt T2D may require additional compromise of the pancreatic β cell function.

Whether IGRP has the same catalytic activity as the hepatic glucose-6-phosphatase encoded by the gene *G6PC1* is a matter of debate (*14*). When IGRP dephosphorylates glucose-6-phosphate, it opposes the action of the β cell glucose sensor glucokinase and hence moderates the glycolytic pathway and eventually glucose-stimulated insulin secretion (*9*). We hypothesize that *G6PC2* is part of the same glucose phosphorylation pathway that includes the glucokinase gene (*GCK*), responsible for the monogenic form of early onset diabetes, MODY2 (*15*), and familial hypoglycemia-hyperinsulinemia phenotypes (*16*). This pathway also includes the glucokinase regulatory protein gene (*GCKR*). Thus, IGRP may antagonize the glucokinase activity in β cells, in a similar manner to the GCKR action as a potent inhibitor of glucokinase activity in hepatocytes (*17*). Recently, SNPs in both *GCK* (rs1799884—30G/A) (*18*) and *GCKR* (rs1260326-P446L) (*2*, *19*, *20*) were found to modulate FPG in European population-based cohorts. To study the interaction of rs560887 with rs1260326-P446L and rs1799884—30G/A, we analyzed their combined effect in the French DESIR cohort and observed an additive effect (interaction $P = 0.56$) of the three SNPs on FPG ($P = 5 \times 10^{-21}$) (Fig. 2). Individuals carrying more than four alleles associated with low FPG (10.7% of the DESIR population) showed a mean 0.24 mmol/l (4.5%) decrease of FPG compared with subjects carrying only one allele (Fig. 2).

G6PC2 rs560887 is associated with significant reductions in both FPG and glycated hemoglobin A1C (HbA1c). There is strong evidence suggesting that even small changes in blood glucose can have an impact on cardiovascular morbidity and mortality (*21*). Meta-analyses have shown that there is a linear relationship, with no threshold, between FPG and CHD risk (*6*), with individuals whose FPG is lower than 5.5 mmol/l

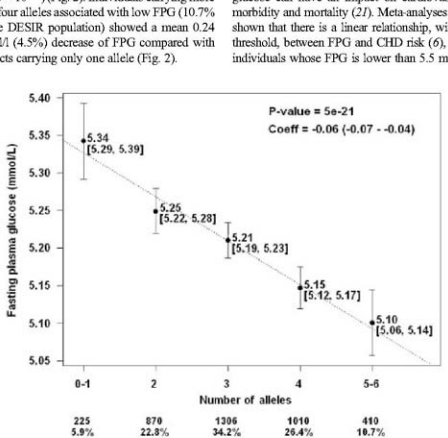


Fig. 2. Combined effects of *G6PC2* rs560887, *GCKR* rs1260326-P446L, and *GCK* rs1799884-30G variants on FPG levels. Data are presented as mean [CI 95%], and the *P* value is for the β coefficient in the linear regression model (including age, sex, and BMI as covariates) of FPG levels on the number of alleles associated with low FPG (*G6PC2* rs560887A allele frequency = 0.30; *GCKR* rs1260326-P446 allele frequency = 0.44; *GCK* rs1799884-30G allele frequency = 0.82).

Table 1. Association of rs560887 with FPG levels in normoglycemic individuals from a GWA study and replication analyses. Association was tested in the additive model, adjusted for age, sex, and BMI. All individuals were normoglycemic (FPG < 6.1 mmol/l). Per A allele effect size was estimated from the regression coefficient β. Overall meta-analysis includes replication populations (DESIR NG, NFBC86, and obese children). DESIR controls are nonobese NG individuals.

			Mean fasting glucose (mmol/l) by rs560887 genotype					
FPG	*N*	MAF	GG	GA	AA	Variance (%)	Per-A allele effect size (mmol/l) [95% CI]	*P*
DESIR controls (stage 1)	654	0.33	5.15 ± 0.33	5.06 ± 0.37	4.93 ± 0.34	3.90	−0.10 [−0.14; −0.06]	4×10^{-7}
DESIR NG (stage 2)	3419	0.30	5.26 ± 0.45	5.19 ± 0.46	5.14 ± 0.47	0.74	−0.06 [−0.08; −0.04]	3×10^{-8}
DESIR cohort (stage 1 and stage 2)	4073	0.30	5.24 ± 0.44	5.17 ± 0.45	5.10 ± 0.46	1.05	−0.07 [−0.09; −0.05]	1×10^{-12}
NFBC86	5073	0.30	5.18 ± 0.40	5.10 ± 0.42	5.05 ± 0.40	1.10	−0.07 [−0.08; −0.05]	6×10^{-16}
Obese children	861	0.28	4.94 ± 0.43	4.88 ± 0.43	4.80 ± 0.47	1.09	−0.07 [−0.11; −0.02]	4×10^{-3}
Overall meta-analysis	9353	0.30	5.18 ± 0.43	5.11 ± 0.44	5.06 ± 0.44	0.91	−0.06 [−0.08; −0.05]	4×10^{-23}

having the lowest mortality rate (*5*, *22*). In contrast, individuals whose FPG lies between 5.6 and 6 mmol/l, although still considered NG, showed a 30% increased risk of heart failure and a 60% increased risk of nephropathy (*23*). In Asians, a slight reduction in FPG from 5.5 mmol/l to 5 mmol/l is associated with a 25% reduction in CHD risk (*24*). Even tight blood glucose control, reflected by HbA1c between 5.1 and 5.6% (normal value < 6.1%), is associated with a 30% higher mortality compared with a lower HbA1c (*25*). In this context, Marz *et al.* previously reported that the glucokinase rs1799884/-30 A allele was associated with T2D and increased risk for coronary artery disease in both diabetics and nondiabetics (*26*). We therefore conclude that the three SNPs identified in our study, which additively modulate FPG by 0.24 mM in the general population, are likely to have a nonnegligible impact on human health.

We found a strong association between SNP rs560887, located in the *G6PC2* gene, and FPG, and this association was replicated in three independent populations (N = 9353). The magnitude of the rs560887 effect on FPG is similar in each cohort [explaining ~1% of FPG variance in each population (heterogeneity test P = 0.95)]. Each copy of the A allele results in a 0.06 mmol/l decrease (overall meta-analysis $P = 4 \times 10^{-23}$). Our findings underscore the role of the glucose phosphorylation pathway in glucose homeostasis in the general population.

References and Notes

1. R. Sladek *et al.*, *Nature* **445**, 881 (2007).
2. R. Saxena *et al.*, *Science* **316**, 1331 (2007); published online 26 April 2007 (10.1126/science.1142358).
3. L. J. Scott *et al.*, *Science* **316**, 1341 (2007); published online 25 April 2007 (10.1126/science.1142382).
4. E. Zeggini *et al.*, *Science* **316**, 1336 (2007); published online 25 April 2007 (10.1126/science.1142364).
5. DECODE Study Group, European Diabetes Epidemiology Group, *Diabetes Care* **26**, 688 (2003).
6. M. Coutinho, H. C. Gerstein, Y. Wang, S. Yusuf, *Diabetes Care* **22**, 233 (1999).
7. Materials and methods are available as supporting material on *Science* Online.
8. J. J. Shieh, C. J. Pan, B. C. Mansfield, J. Y. Chou, *FEBS Lett.* **562**, 160 (2004).
9. Y. Wang *et al.*, *Diabetes* **57**, 133 (2008).
10. Y. Wang *et al.*, *Diabetologia* **50**, 774 (2007).
11. M. Pertea, X. Lin, S. L. Salzberg, *Nucleic Acids Res.* **29**, 1185 (2001).
12. S. Cauchi *et al.*, *Diabetes* **55**, 3189 (2006).
13. R. J. Loos *et al.*, *Diabetes* **56**, 1943 (2007).
14. C. C. Martin *et al.*, *J. Biol. Chem.* **276**, 25197 (2001).
15. P. Froguel *et al.*, *Nature* **356**, 162 (1992).
16. A. L. Cuesta-Munoz *et al.*, *Diabetes* **53**, 2164 (2004).
17. E. Van Schaftingen, M. Detheux, M. Veiga da Cunha, *FASEB J.* **8**, 414 (1994).
18. M. Vaxillaire *et al.*, *Diabetes* **57**, 244 (2008).
19. T. Sparso *et al.*, *Diabetologia* **51**, 70 (2008).
20. M. Orho-Melander *et al.*, *Diabetologia* **50** (S1), S40 (2007).
21. C. Panzer, M. S. Lauer, A. Brieke, E. Blackstone, B. Haegwert, *Diabetes* **51**, 803 (2002).
22. B. Balkau, S. Bertrais, P. Ducimetiere, E. Eschwege, *Diabetes Care* **22**, 696 (1999).
23. K. T. Khaw *et al.*, *BMJ* **322**, 15 (2001).
24. C. M. Lawes *et al.*, *Diabetes Care* **27**, 2836 (2004).
25. E. B. Levitan *et al.*, *Diabetologia* **51**, 267 (2008).
26. W. Marz *et al.*, *Circulation* **109**, 2844 (2004).

27. We thank O. Lantieri for comments on the manuscript; V. Vatin, M. Dewilder, F. Allegaert, S. Gaget, S. Galfina, J.-C. Chivre, and J. Delplanque for technical assistance; and L. Peltonen for providing DNA samples. The study was supported by ALFEDIAM-Servier, ANR-06 PHYSIO-037-02, EURODIA ISHM-CT-2006-518153, Genome Canada, Genome Quebec, KU-Leuven GOA/2004/11, Juvenile Diabetes Research Foundation 1-2006-182, Imperial College London, and the Academy of Finland. NFBC was supported by the European Commission, contract no. Q LG1-CT-2000-01643, and the Academy of Finland.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1156849/DC1
Materials and Methods
Figs. S1 and S2
Tables S1 to S3
References

21 February 2008; accepted 3 April 2008
Published online 1 May 2008;
10.1126/science.1156849
Include this information when citing this paper.

The Serine Protease TMPRSS6 Is Required to Sense Iron Deficiency

Xin Du,[1] Ellen She,[1] Terri Gelbart,[2] Jaroslav Truksa,[2] Pauline Lee,[2] Yu Xia,[1] Kevin Khovananth,[1] Suzanne Mudd,[1] Navjiwan Mann,[1] Eva Marie Y. Moresco,[1] Ernest Beutler,[2] Bruce Beutler[1]*

Hepcidin, a liver-derived protein that restricts enteric iron absorption, is the key regulator of body iron content. Several proteins induce expression of the hepcidin-encoding gene *Hamp* in response to infection or high levels of iron. However, mechanism(s) of *Hamp* suppression during iron depletion are poorly understood. We describe *mask*: a recessive, chemically induced mutant mouse phenotype, characterized by progressive loss of body (but not facial) hair and microcytic anemia. The *mask* phenotype results from reduced absorption of dietary iron caused by high levels of hepcidin and is due to a splicing defect in the transmembrane serine protease 6 gene *Tmprss6*. Overexpression of normal TMPRSS6 protein suppresses activation of the *Hamp* promoter, and the TMPRSS6 cytoplasmic domain mediates *Hamp* suppression via proximal promoter element(s). TMPRSS6 is an essential component of a pathway that detects iron deficiency and blocks *Hamp* transcription, permitting enhanced dietary iron absorption.

Iron is an essential cofactor for a host of metabolic reactions, necessary for the proper functioning of the oxygen-carrying protein hemoglobin and proteins of the electron transport chain. Though required for life, iron is also a toxic element when present in excess, and body iron content is normally maintained within narrow limits, controlled by the efficiency of intestinal iron absorption. Hepcidin, a key regulator of iron absorption (*1*), binds to ferroportin, the channel for cellular iron efflux, and causes internalization and proteolysis of the channel, preventing release of iron from macrophages or intestinal cells (*2*) into the plasma. In this manner, hepcidin lowers plasma iron levels, and chronic elevation of hepcidin levels causes systemic iron deficiency (*3*). Conversely, hepcidin deficiency causes systemic iron overload (*1*). Several proteins—including HFE (hemochromatosis protein) (*4*), transferrin receptor 2 (*5*), hemojuvelin (*6*), and the transcription factor Smad4 (*7*)—regulate body iron levels by promoting expression of *Hamp*, the gene encoding hepcidin. However, the components of a pathway to down-regulate *Hamp* in response to iron deficiency (*8*) remain poorly understood.

We now report the discovery of a protein with nonredundant function in *Hamp* suppression, required to permit normal uptake of enteric iron in response to iron depletion and capable of countering all known *Hamp*-activating stimuli. A phenotype observed among third generation (G3) mice (C57BL/6J background) homozygous for mutations induced by *N*-ethyl-*N*-nitrosourea (*9*), termed *mask*, was characterized by gradual loss of body but not facial hair, leading to almost complete nudity of the trunk within about 4 weeks (Fig. 1A). The *mask* homozygotes are slightly smaller than their heterozygous littermates, and adult female homozygotes are infertile. Male homozygotes retain fertility, as do heterozygotes of both sexes. Mutant mice maintained on a standard laboratory diet manifested microcytic anemia, low plasma iron levels (Fig. 1B), and depleted iron stores (fig. S1). There was no evidence of occult blood loss. We therefore suspected that intestinal iron absorption might be abnormal and measured the retention of ferrous ^{59}Fe after intragastric administration to *mask* homozygotes and wild-type controls receiving a standard laboratory diet. The *mask* mice retained slightly more iron than wild-type animals (Fig. 1C). However, because *mask* mice were anemic and iron-deficient (conditions that favor iron absorption), we repeated the experiment using *mask* homozygotes and controls that had been maintained on an iron-deficient diet for 4 weeks. The relationship between *mask* homozygotes and controls was now reversed (Fig. 1D): Iron-deprived but nonanemic control animals showed about a sixfold increase in efficiency of iron absorption, whereas anemic *mask* homozygotes

[1]Department of Genetics, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA. [2]Department of Molecular and Experimental Medicine, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA.

*To whom correspondence should be addressed. E-mail: bruce@scripps.edu

showed only a 3.5-fold increase in comparison with animals maintained on a standard laboratory diet.

In addition, whereas iron deprivation led to suppression of *Hamp* transcript levels in the livers of wild-type mice, high *Hamp* transcript levels were observed in the livers of anemic *mask* homozygotes (Fig. 1E). This outcome was consistent with insensitivity to low iron stores and consequent failure to suppress *Hamp* expression. The aberrant up-regulation of *Hamp* transcription occurred despite anemia, which (as distinct from iron deficiency alone) normally lowers *Hamp* transcript levels (*8*). Both fertility and hair growth (Fig. 1, F and G) were consistently restored when *mask* homozygotes were maintained on an iron-enriched diet (2% carbonyl iron). Hence, the visible phenotype that led to the identification of *mask* was a consequence of iron deficiency.

The mutation responsible for the *mask* phenotype was mapped to distal chromosome 15 and then further confined to a region 1.3 megabase pairs in size (*10*). A mutation was identified in *Tmprss6*, encoding a transmembrane serine protease of unknown function. TMPRSS6, also known as matriptase-2 (*11*), is a type II plasma membrane protein that is 811 residues in length. Expression anatomy databases indicate that the *Tmprss6* gene is expressed in a limited number of tissues, including the liver and the olfactory and vomeronasal epithelium in the mouse [(*12*) and fig. S2]. In both humans and mice, the major site of TMPRSS6 expression is the liver. TMPRSS6 includes a C-terminal trypsin-like serine protease domain (shown to be extracellular), three class A low-density lipoprotein receptor domains, two CUB domains [similar to a domain represented in bone morphogenetic protein 1 (BMP1) as well as C1R and C1S proteins], and a membrane-proximal SEA domain (Fig. 2B). The *mask* mutation, an A→G transition, eliminates a splice acceptor site upstream from exon 15 (Fig. 2A and fig. S3). Complementary DNA sequencing showed that two abnormal splice products result from this mutation (GenBank accession numbers EU190436 and EU190437), both encoding products that would lack the proteolytic domain (Fig. 2B). No normal mRNA product was detectable in homozygous mutants (fig. S4). We confirmed that the *Tmprss6* mutation is responsible for the *mask* phenotype by bacterial artificial chromosome (BAC) transgenesis (figs. S5 and S6). Introduction of a 197-kb piece of genomic DNA containing the intact *Tmprss6* gene and its 5′ and 3′ flanking elements rescued the hair-loss phenotype in four of four homozygous mutant mice and rescued anemia and iron deficiency in three of four transgenic mice (Fig. 3) (*10*).

Hamp is up-regulated in response to the inflammatory cytokines interleukin-6 (IL-6) (*13*) and IL-1 (*14*), which may collectively mediate the anemia of chronic disease (*15*). It is also up-regulated by the bone morphogenetic proteins [e.g., BMP2, BMP4, and BMP9 (*7*, *16*)] and by hemojuvelin (*17*). Our observation of low iron levels coupled with high *Hamp* transcript levels in *mask* mutants suggested that TMPRSS6 might negatively regulate *Hamp* activation to promote iron uptake. We hypothesized that the intact TMPRSS6 protein, if overexpressed, might prevent induced up-regulation of the *Hamp* promoter. We tested this hypothesis by cotransfecting HepG2 cells with a normalized *Hamp* reporter construct (*18*, *19*) and cDNAs encoding either wild-type or *mask* mutant versions of TMPRSS6. Either a cotransfected hemojuvelin- or Smad1-encoding sequence—or stimulation with IL-1α, IL-6, BMP2, BMP4, or BMP9—was used to induce the *Hamp* promoter (Fig. 4A and fig. S7). TMPRSS6 cotransfection strongly inhibited *Hamp* reporter gene activation by each stimulus, whereas the *mask* mutant version of TMPRSS6 showed (at most) a modest inhibitory effect. However, the mutant TMPRSS6 appeared to retain some suppressing activity, particularly with respect to its ability to block *Hamp* promoter responses to

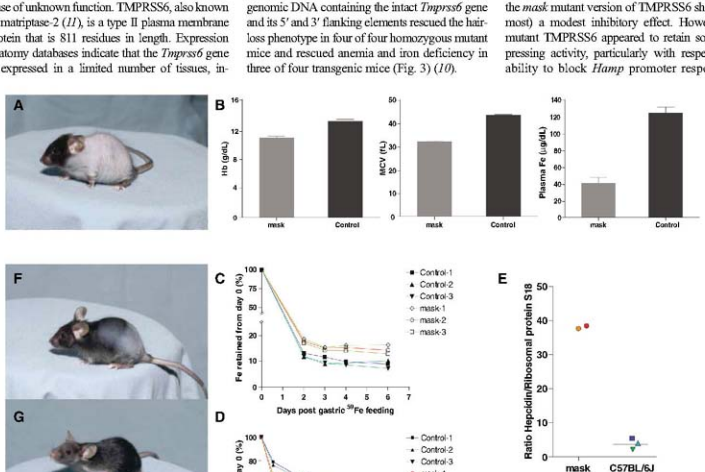


Fig. 1. The *mask* phenotype. **(A)** Appearance of *mask* mice at 4 weeks of age. **(B)** Anemia and iron deficiency in *mask* homozygotes 6 to 8 weeks of age. Hemoglobin (Hb), mean red blood cell volume (MCV), and plasma iron levels in *mask* homozygote and control mice are shown. $n = 4$ mice for each column. Error bars represent 1 SD. **(C** and **D)** Retention of radioactive iron administered intragastrically. In (C), mice were maintained on a normal diet before instillation of iron. In (D), mice were iron-depleted for 1 month before instillation of iron. **(E)** After measuring retention of ^{59}Fe in iron-depleted mice, hepcidin mRNA levels were measured in the liver of the same five animals as in (D), normalized to mRNA levels of ribosomal protein S18. Horizontal lines indicate mean values. Color coding is preserved for individual mice. Regrowth of hair in the mouse shown in **(A)**, 1 week **(F)** and 2 weeks **(G)** after initiation of a high-iron diet.

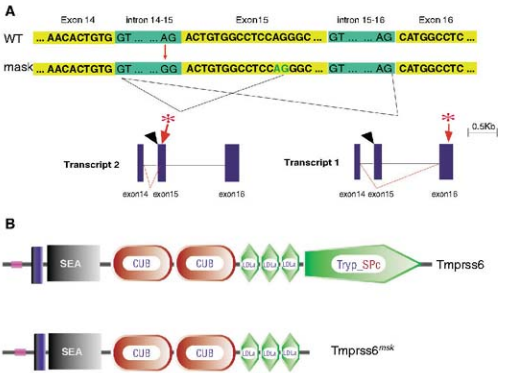


Fig. 2. The *mask* mutation. (**A**) Schematic illustration of the transcripts resulting from the mutated *Tmprss6*^{msk} allele. The mutation site, at the acceptor splice junction upstream from exon 15, is indicated by the black arrowhead; a premature stop codon is predicted as indicated by the asterisk. WT, wild type. (**B**) Illustration of the coding change that results from the mutation [Simple Modular Architecture Research Tool (SMART) diagram]. Tryp_SPc, trypsin-like serine protease; LDLa, low-density lipoprotein receptor domain class A.

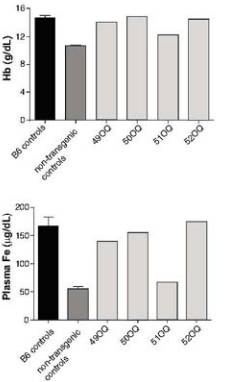


Fig. 3. Rescue of the *mask* phenotype by BAC transgenesis. Four founders (designated 49 OQ to 52 OQ) were produced by microinjecting a BAC clone bearing the WT *Tmprss6* sequence into single-cell embryos homozygous for the *mask* mutation. Hemoglobin and serum iron levels in each of the founders, in blood sampled at 4 weeks of age, are shown. B6 controls, age-matched C57BL/6J mice, n = 4 mice; non-transgenic controls, *Tmprss6*$^{msk/msk}$ littermates that lacked the transgene, n = 4 mice. Error bars indicate 1 SEM.

IL-1α and IL-6. Smad1-induced *Hamp* promoter responses were also blocked by normal (but not truncated) TMPRSS6 (fig. S7).

We hypothesized that TMPRSS6 participates in a transmembrane signaling pathway triggered by iron deficiency and independent of the known *Hamp* activation pathways. This putative pathway would ultimately interact with suppressive element(s) in the *Hamp* promoter, inhibiting expression in response to activating stimuli. To determine whether the proteolytic activity of TMPRSS6 was essential to suppression of *Hamp* promoter activity, we mutated the serine residue of the catalytic triad to alanine, and the protease-dead mutant was expressed in HepG2 cells together with activating stimuli. The point mutation produced an effect similar to that of the *mask* truncation, indicating that proteolysis by TMPRSS6 is integral to the *Hamp* suppression mechanism (Fig. 4B). To address whether TMPRSS6 might signal via its own cytoplasmic domain, we designed two additional constructs, termed delN and delC. delN retained the entire TMPRSS6 ectodomain and membrane-spanning domain, but the cytoplasmic domain was replaced with a green fluorescent protein (GFP)–encoding sequence. delC consisted of a TMPRSS6 cytoplasmic domain and membrane-spanning domain, but the ectodomain was replaced with a GFP-encoding sequence. Both delN and delC were membrane-expressed proteins, as assessed by confocal fluorescence microscopy (Fig. 4C). When overexpressed, both delN and delC suppressed *Hamp* promoter activity more strongly than the wild-type construct (Fig. 4D).

Interpreting these results in light of the fact that an endogenous human *TMPRSS6* transcript is expressed by HepG2 cells, we infer that (i) TMPRSS6 is in a signaling pathway that suppresses *Hamp* expression; (ii) the TMPRSS6 ectodomain represses this signal, just as many receptor ectodomains enforce receptor silence in the absence of a specific ligand (*20*, *21*); and (iii) the cytoplasmic domain of TMPRSS6 is integral to signal transduction, because when overexpressed in the absence of the repressive ectodomain, it autonomously drives a *Hamp* suppression signal. Overexpression of the wild-type protein or the delN construct suppresses *Hamp* expression, and the TMPRSS6 ectodomain is only capable of signaling if it is capable of proteolysis. A model that would account for these observations is presented in fig. S8.

To analyze *Hamp* promoter responses to activating and inhibitory stimuli, we tested a series of deletion constructs for activation by IL-6 administered to HepG2 cells in the presence or absence of TMPRSS6 overexpression. As little as 140 base pairs (bp) of promoter sequence (upstream from the cap site) supported the TMPRSS6-mediated blockade of reporter gene activation. The full activating effect of IL-6 required roughly 260 bp of promoter sequence (Fig. 4E). A proximal repressor element in the *Hamp* promoter thus mediates inhibition by a TMPRSS6-dependent signal, although it has yet to be determined whether other repressor elements are active within more distal portions of the promoter.

The *Tmprss6* locus has not been targeted in mice. TMPRSS6 is known to cleave extracellular matrix proteins (*11*). However, as demonstrated by the *mask* phenotype, a major function of this protein is in iron regulation. TMPRSS6-mediated *Hamp* suppression permits adequate uptake of iron from dietary sources. Without *Hamp* suppression, severe iron deficiency supervenes despite "normal" dietary iron intake. Hepcidin-mediated restriction of iron absorption operates effectively over a range of iron concentrations, encompassing that found in the diet of laboratory mice. However, hepcidin cannot restrict the absorption of supraphysiologic concentrations of iron, which occurs through pathway(s) that permit uptake despite systemic iron overload (*22*). Thus, *mask* mutants exhibit dramatic iron deficiency when fed a standard diet but show phenotypic normalization in response to oral administration of 2% carbonyl iron (Fig. 1, F and G, and fig. S9). The mechanism of hair loss in *mask* homozygotes remains to be established. It is of note that severe iron deficiency is associated with hair loss in humans (*23*) as it is in mice (*24*).

It has been suggested that the proteins hypoxia-inducible factor 1α (Hif-1α) and growth and differentiation factor 15 (GDF15) participate in the physiologic response to body iron attenuation. Hif-1α is a transcription factor regulated by the iron-dependent prolyl-4-hydroxylase (PHD2) (*25*)

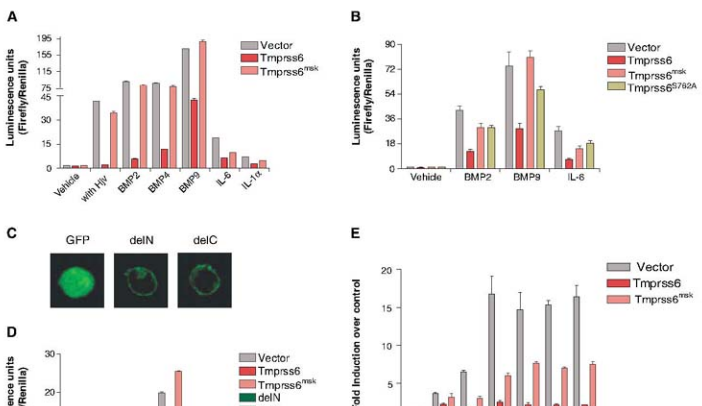


Fig. 4. Effect of WT or *mask* mutant TMPRSS6 expression on *Hamp* promoter responses. HepG2 cells were transfected with a full-length *Hamp* promoter reporter construct in all experiments shown in (A), (B), and (D) and with a normalizing renilla luciferase construct. Tmprss6, full-length construct encoding the native protein; Tmprss6^{msk}, the truncated mutant sequence expressed by *mask* mice; Tmprss6^{S762A}, protease-dead Tmprss6 mutant (S762A, Ser^{762}→Ala^{762}). delN: GFP cytoplasmic domain, and TMPRSS6 transmembrane and ectodomain; delC: GFP ectodomain, and TMPRSS6 transmembrane and cytoplasmic domains. **(A)** The effect of Tmprss6 and Tmprss6^{msk} on *Hamp* promoter activity in cells cotransfected with a hemojuvelin (Hjv)–encoding vector or stimulated with BMP2, BMP4, BMP9, IL-6, or IL-1α. **(B)** Comparison of the *Hamp* promoter-suppressing effects of Tmprss6, Tmprss6^{msk}, and Tmprss6^{S762A} in cells subjected to the stimuli indicated. **(C)** Localization of delN and delC constructs in HepG2 cells by confocal fluorescence microscopy. **(D)** Comparison of the *Hamp* promoter-suppressing effects of Tmprss6, Tmprss6^{msk}, delN, and delC constructs in cells subjected to the stimuli indicated. **(E)** Response of *Hamp* promoter deletion constructs to induction by IL-6 (length indicated with respect to the cap site) and the relative suppression of responses by cotransfection with Tmprss6 or Tmprss6^{msk} constructs. In all experiments [(A) to (E)], duplicate transfections were performed. Incubations with cytokines were performed for 16 hours before luminescence was read to estimate hepcidin promoter activity. Error bars indicate 1 SEM.

and responds both to low cellular iron concentration and to hypoxia (*26*–*28*). GDF15 is a circulating transforming growth factor–β superfamily member that suppresses *Hamp* expression, albeit under conditions of iron overload, in the context of thalassemia (*29*). Either of these proteins might be elements of the TMPRSS6-dependent pathway as we have described it. However, the genetic evidence presented here demonstrates that neither Hif-1α nor GDF15 supplants the function of TMPRSS6 under standard feeding conditions. We also note that the TMPRSS6-dependent pathway predominates over all known *Hamp*-activating pathways, because overexpression of TMPRSS6 can nullify—in part or in full—the hepcidin-inducing effects of hemojuvelin, BMP2, BMP4, BMP9, Smad1, IL-1, and IL-6.

Because overexpression of TMPRSS6 blocks activation of diverse pathways for up-regulation of hepcidin, we anticipate that overexpression or dysregulation of TMPRSS6 might cause iron overload, whereas mutational inactivation of TMPRSS6 (as in *mask* mice) would lower body iron content. From a therapeutic standpoint, TMPRSS6 activation might block cytokine-mediated anemia during chronic disease, and selective antagonism of TMPRSS6 might be used in the treatment of systemic iron overload (hemochromatosis). With regard to its olfactory and vomeronasal expression, we suggest that in addition to signaling hepatocytes of iron deficiency, TMPRSS6 may alert olfactory neurons and thereby engender iron-seeking behavior under conditions of iron starvation. Further work will be required to determine whether this is the case.

References and Notes

1. E. Nemeth, T. Ganz, *Annu. Rev. Nutr.* **26**, 323 (2006).
2. E. Nemeth *et al.*, *Science* **306**, 2090 (2004).
3. G. Nicolas *et al.*, *Proc. Natl. Acad. Sci. U.S.A* **99**, 4596 (2002).
4. K. A. Ahmad *et al.*, *Blood Cells Mol. Dis.* **29**, 361 (2002).
5. H. Kawabata *et al.*, *Blood* **105**, 376 (2005).
6. G. Papanikolaou *et al.*, *Nat. Genet.* **36**, 77 (2004).
7. R. H. Wang *et al.*, *Cell Metab.* **2**, 399 (2005).
8. G. Nicolas *et al.*, *J. Clin. Invest.* **110**, 1037 (2002).
9. Materials and methods are available as supporting material on *Science* Online.
10. See the supporting material on *Science* Online.
11. G. Velasco, S. Cal, V. Quesada, L. M. Sanchez, C. Lopez-Otin, *J. Biol. Chem.* **277**, 37637 (2002).
12. Genomics Institute of the Novartis Research Foundation (GNF) SymAtlas, http://symatlas.gnf.org/SymAtlas/.
13. E. Nemeth *et al.*, *Blood* **101**, 2461 (2003).
14. P. Lee, H. Peng, T. Gelbart, L. Wang, E. Beutler, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 1906 (2005).
15. C. N. Roy, N. C. Andrews, *Curr. Opin. Hematol.* **12**, 107 (2005).
16. J. Truksa, H. Peng, P. Lee, E. Beutler, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 10289 (2006).
17. J. L. Babitt *et al.*, *Nat. Genet.* **38**, 531 (2006).
18. J. Truksa, P. Lee, H. Peng, J. Flanagan, E. Beutler, *Blood* **110**, 3436 (2007).
19. J. Truksa, H. Peng, P. Lee, E. Beutler, *Br. J. Haematol.* **139**, 138 (2007).
20. F. Bazzoni, B. Beutler, *N. Engl. J. Med.* **334**, 1717 (1996).
21. S. S. Watowich, D. J. Hilton, H. F. Lodish, *Mol. Cell. Biol.* **14**, 3535 (1994).
22. D. Gitlin, A. Cruchaud, *J. Clin. Invest.* **41**, 344 (1962).
23. C. Delache *et al.*, *Eur. J. Dermatol.* **17**, 507 (2007).

24. E. S. Russell, E. C. McFarland, E. L. Kent, *Transplant. Proc.* **2**, 144 (1970).
25. L. A. McNeill *et al.*, *Mol. Biosyst.* **1**, 321 (2005).
26. C. Peyssonnaux *et al.*, *J. Clin. Invest.* **117**, 1926 (2007).
27. D. Yoon *et al.*, *J. Biol. Chem.* **281**, 25703 (2006).
28. E. Beutler, *Science* **306**, 2051 (2004).
29. T. Tanno *et al.*, *Nat. Med.* **13**, 1096 (2007).

30. This work was supported by NIH (grant numbers AI054523 and DK53305-09) and by the Stein Endowment Fund.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1157121/DC1
Materials and Methods
SOM Text
Figs. S1 to S9
References

28 February 2008; accepted 17 April 2008
Published online 1 May 2008;
10.1126/science.1157121
Include this information when citing this paper.

The Right and the Good: Distributive Justice and Neural Encoding of Equity and Efficiency

Ming Hsu,[1]* Cédric Anen,[2]* Steven R. Quartz[2]†

Distributive justice concerns how individuals and societies distribute benefits and burdens in a just or moral manner. We combined distribution choices with functional magnetic resonance imaging to investigate the central problem of distributive justice: the trade-off between equity and efficiency. We found that the putamen responds to efficiency, whereas the insula encodes inequity, and the caudate/septal subgenual region encodes a unified measure of efficiency and inequity (utility). Notably, individual differences in inequity aversion correlate with activity in inequity and utility regions. Against utilitarianism, our results support the deontological intuition that a sense of fairness is fundamental to distributive justice but, as suggested by moral sentimentalists, is rooted in emotional processing. More generally, emotional responses related to norm violations may underlie individual differences in equity considerations and adherence to ethical rules.

I magine driving a truck with 100 kg of food to a famine-stricken region. The time it would take you to deliver food to everyone would cause 20 kg of food to spoil. If you delivered food to only half the population, you would lose only 5 kg. Do you deliver the food to only half the population to maximize the total amount of food, or do you sacrifice 15 kg to help everyone and achieve a more equitable distribution?

This dilemma illustrates the core issues of distributive justice, which involves trade-offs between considerations that are at once compelling but that cannot be simultaneously satisfied. More generally, distributive justice concerns how individuals and societies allocate benefits and burdens in a just or moral manner, and it is central to social choice theory, moral psychology, and welfare economics (*1*–*3*). Despite the long history of work on distributive justice, however, its psychological and neural underpinnings remain poorly understood, much of it centering on two long-standing debates.

The first debate concerns the role of equity and fairness: Is it more just to maximize some overall good (such as well-being) independently of its distribution, or must its distribution satisfy certain criteria (such as equity), even if it results in less overall well-being? Utilitarian theories of justice, exemplified by Mill and Harsanyi, maximize the good, or efficiency. In its simplest form, this involves maximizing the sum of individual utilities, irrespective of equity (*4*). In contrast, deontological theories of distributive justice maintain that the right (e.g., equity) is prior to the good and that an action can maximize the good and yet be morally wrong if it violates a deontological principle, such as a rule, right, or duty. Contemporary proponents of deontological theories, most notably Rawls, observe the near universality of fairness norms and argue that this sense of fairness underlies institutions and society as a whole, thereby generating the notion of "justice as fairness" (*5*).

A second debate concerns the involvement of emotion in distributive justice. A prominent cognitivist tradition, including such philosophers as Plato and Kant, emphasizes the role of reason in resolving the trade-off between the right and the good, as do many contemporary thinkers including Rawls and Harsanyi (*4*, *5*). In psychology, a prominent cognitivist view suggests that a sense of justice emerged as a developmental consequence of formal and abstract cognition (*2*). An alternative tradition, including moral sentimentalists such as David Hume and Adam Smith, argues that distributive justice is rooted in emotions, such as sympathy and empathy (*6*, *7*).

Although these debates remain unresolved, recent works in related fields—including moral judgment and economics—provide converging evidence of the interplay between emotion and cognition (*8*, *9*), as well as the importance of fairness (*10*–*12*), in individual and social decision-making. Based on these findings, we hypothesized that distinct neural substrates may underlie the representation of equity and efficiency. First, we hypothesized that reward regions, such as the striatum, would be involved in encoding utility and efficiency. A wide variety of decision-making studies indicate the involvement of dopaminergic regions, such as the striatum, in reward computation and reward learning (*13*, *14*), including indirect rewards such as charitable giving and punishment of free-riders in public-goods games (*15*, *16*). More recent evidence has implicated nearby paralimbic regions, especially the septal-subgenual area, in altruism and social attachment (*17*, *18*).

Second, we hypothesized that emotional systems, particularly the insular cortex, would be involved in the encoding of inequity, as recent work has demonstrated the important role of the insular cortex in fairness and empathy (*9*, *19*, *20*). We also note the deep connection that exists in economic theory between decision-making under uncertainty and the measurement of inequity (*21*–*23*). This connection is of particular relevance in light of growing evidence that the insular cortex is involved in risky decisions and risk perception (*24*, *25*), as well as a possible separation in the encoding of reward and risk (*25*). Finally, we speculated that differential activation of reward and emotional regions may reflect their trade-off between efficiency and equity and that such differential activation may correlate with individual differences in these decisions.

To investigate the neural foundations of distributive justice, we employed a novel distribution task in conjunction with functional magnetic resonance imaging (fMRI). During functional brain imaging, 26 adult participants (nine males, mean age of 39.2 years, age range of 29 to 55 years) made decisions about how to allocate money to a group of children living in an orphanage in northern Uganda (*23*). Each group of three children was endowed by the experimenters with $5, the monetary equivalent of 24 meals per child. We denominated allocations in meals to give participants an approximation of the purchasing power of the monies being donated (*23*). In each trial, participants decided whether varying allocations of money, denominated in meals, would be taken away from either of two groups of children; the participant's choice was to decide from whom to take. Participants donated $87 on average (for a total of $2279) to the charity (*23*).

This design allowed us to parametrically vary the relative efficiency and equity of the allocations, providing a quantitative framework to evaluate participants' choices (fig. S3). Specifically, we used an inequity-aversion model in which individuals trade off between equity and efficiency (*26*). The additive nature of the model, together

[1]Beckman Institute for Advanced Science and Technology and Department of Economics, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. [2]Social Cognitive Neuroscience Laboratory, Division of Humanities and Social Sciences 228-77, California Institute of Technology, Pasadena, CA 91125, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: steve@hss.caltech.edu

with the design of the experiment in which equity and efficiency were varied independently, allowed us to create orthogonal regressors and explore the possibility of separate neural encoding of these two variables (*23*). The dynamic nature of our task partitioned the temporal ordering of the trials (Fig. 1 and movies S1 and S2), allowing for an event-related fMRI analysis and a search for possible differential neural contributions at various decision-making stages, as time is a crucial variable in brain systems involved in reward and learning (*13*) and models of decision-making (*27*).

Behaviorally, the group inequity-aversion parameter estimate was $\alpha = 6.96 \pm 1.08$ (*23*). In addition, individual inequity-aversion estimates showed substantial variation (fig. S4 and table S3) (*23*), which further allowed us to use the estimated individual inequity-aversion attitude as a

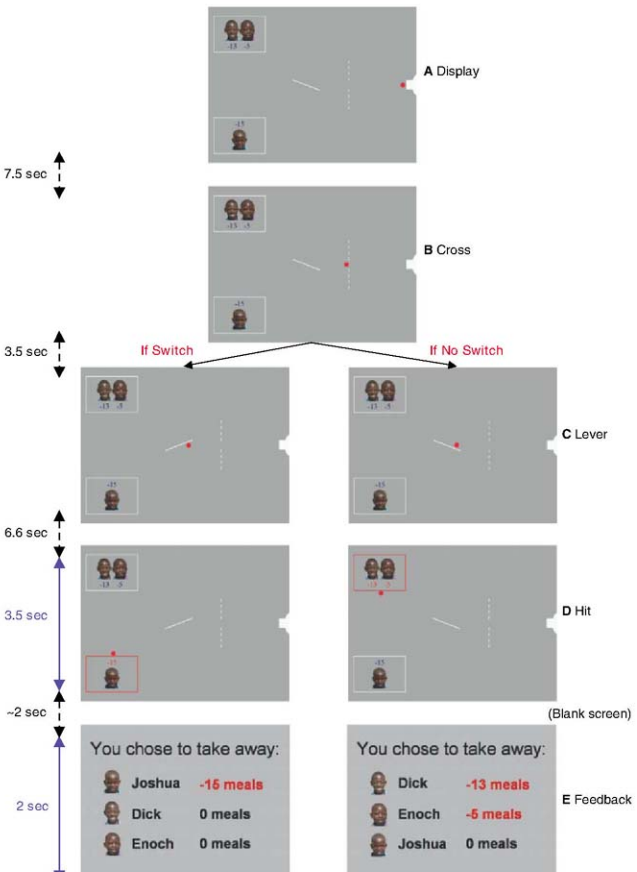


Fig. 1. Timeline and animation stills of experimental design. **(A)** Display: A projectile begins moving across the screen toward two groups of children. The number of meals each child can potentially lose (from an initial endowment of 24) is given next to the picture of the child. The position of the lever in the middle of the screen denotes the default group of children that will lose the meals. **(B)** Cross: After the projectile crosses the dotted line, the participant may switch the lever (Switch) to direct the projectile toward the other group of children. The participant may only switch the lever once and has 3.5 s to do so. **(C)** Lever: Once the projectile hits the lever, the participant can no longer switch it. The projectile continues to move toward the group of children from whom the participant has chosen/allowed the meals to be taken away. **(D)** Hit: The projectile touches the box surrounding the pictures; the box changes color and remains highlighted for 3.5 s. **(E)** Feedback: After a blank screen of random duration (uniformly distributed from 1 to 3 s), a feedback screen (2 s) informs the participant how many meals each child received. Subsequent trials are separated by a blank screen of random duration (uniformly distributed from 5 to 7 s).

between-participant measure in the neuroimaging data analysis.

Functional imaging results were analyzed using standard regression techniques (*23*). An event-related design was used where regressors were included for the various events of the trials (Fig. 1). Interaction terms corresponding to efficiency and equity, or utility, were added as parametric regressors (*23*).

We first searched for regions that respond to both efficiency and inequity in the form of the hypothesized utility function (*23*). Figure 2A shows activation of a region overlapping with the caudate head and the septal-subgenual area, with respect to the marginal utility of participants' choices (ΔU) during the "Hit" event only (*23*). The activation was driven by both marginal efficiency (ΔM) and marginal inequity (ΔG) and is the only region that survives at the uncorrected threshold of $P < 0.001$ (Fig. 2A). Furthermore, because ΔU was calculated with the group-level inequity-aversion parameter α, the inequity-aversion model predicts that the coefficients would be negatively correlated with the individual inequity-aversion estimates. That is, individuals with higher neural responses to inequity would reject the inequitable allocation in favor of one that is more equitable. Figure 2B shows that this is indeed the case (Spearman $\rho = -0.419$, $P < 0.02$, two-tailed) (*28*).

Next, we looked for the hypothesized separation in neural regions encoding efficiency and equity. Figure 3A shows bilateral putamen activation with respect to M_C, the efficiency of chosen allocations, during the "Display" event only and again is the only region that survives at the uncorrected threshold of $P < 0.001$ (*23*). Unlike the ΔU region, however, the putamen is correlated only with efficiency rather than inequity (Fig. 3B) and is also not correlated with individual inequity-aversion parameters (table S3).

In contrast, we found that activity in the bilateral insular cortex is significantly correlated with ΔG during the Display (and also the "Switch") event (Fig. 4A and fig. S4). The amount of inequity reduced by the participant's choice is therefore a monotonic function of insular activation. Furthermore, activity in the insula is not significantly correlated with measures of efficiency (Fig. 4B). Strikingly, we found that the individual β values of ΔG are significantly negatively correlated with individual inequity-aversion parameters (Fig. 4C). Therefore, as in the ultimatum game (UG) (*9*), high insula activity is associated with passing over the inequitable allocation and choosing the equitable allocation [the allocation of (0, 0) in the UG] (fig. S7 and table S4). This also supports the more general proposed role of insula in norm violation (*29*) and the idea that individual differences arise from differing sensitivity to inequity norms. That is, participants who receive strong negative affective signals may be more sensitive to violating fairness norms and hence adhere more to deontological norm following, whereas those who do not are influenced primarily by efficiency. As persistent violation of social norms is symptomatic of conduct disorder, our finding of an affective basis to norm following may also have important implications for understanding the disruption in norm following in clinical populations (*30*).

Investigations of distributive justice reach back to the beginning of philosophy and have important social, political, and economic implications (*23*). Our results show how the brain encodes two considerations central to the distributive justice calculus and shed light on the cognitivist/sentimentalist debate regarding the psychological underpinnings of distributive justice. Specifically, the dissociation between the inequity regions on the one hand and the efficiency regions on the other supports the inequity-aversion model we employ, in that individual differences in choice behavior arise from

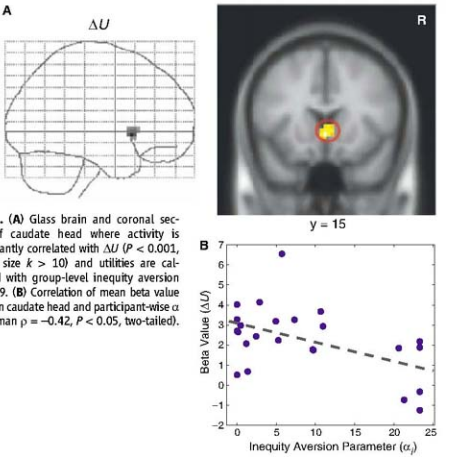


Fig. 2. **(A)** Glass brain and coronal section of caudate head where activity is significantly correlated with ΔU ($P < 0.001$, cluster size $k > 10$) and utilities are calculated with group-level inequity aversion $\alpha = 6.9$. **(B)** Correlation of mean beta value of ΔU in caudate head and participant-wise α (Spearman $\rho = -0.42$, $P < 0.05$, two-tailed).

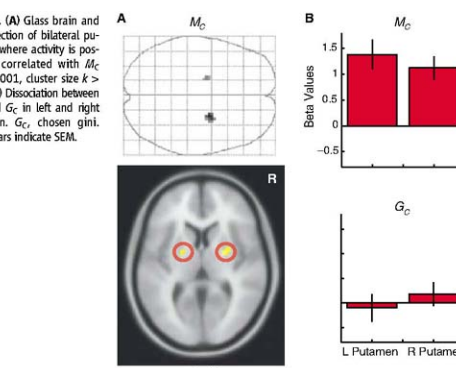


Fig. 3. **(A)** Glass brain and axial section of bilateral putamen where activity is positively correlated with M_C ($P < 0.001$, cluster size $k > 10$). **(B)** Dissociation between M_C and G_C in left and right putamen. G_C, chosen gini. Error bars indicate SEM.

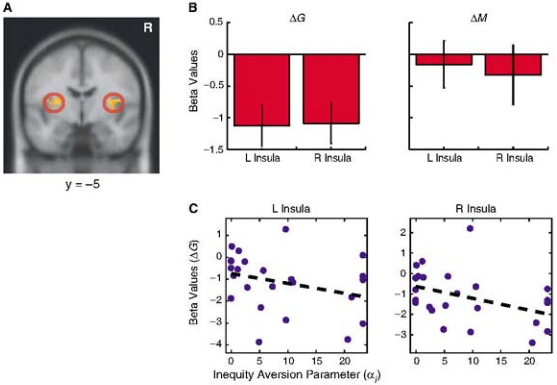


Fig. 4. **(A)** Activation in the bilateral insula during the Display event is negatively correlated with inequity measure ΔG ($P < 0.002$, cluster size $k > 10$). **(B)** Dissociation between ΔG and ΔM in the insula. Error bars indicate SEM. **(C)** Correlation of mean beta value in the insula and participant-wise α (right insula: Spearman $\rho = -0.41$, $P < 0.05$, two-tailed; left insula: $\rho = -0.36$, $P < 0.1$, two-tailed).

participants placing different weights upon inequity, as opposed to efficiency. In addition, we note the substantial degree to which activations reported above are neuroanatomically and temporally distributed across the relevant events of interest (tables S5 to S14). That is, M_C is activated in the putamen only during Display, ΔG is activated in the insula during Display and Switch, and ΔU is activated in the caudate/septal region during Hit. The anatomical separation implies that computation of the hypothesized inequity-aversion utility is distributed in the brain in much the same way that expected reward and risk are proposed to be distributed (*25*). In addition, the separation suggests a role for the region around the caudate head/septal-subgenual system in integrating multiple values into an social evaluatory signal, which is consistent with previous studies implicating this region in social attachment, trust, and charitable givings (*16*–*18*). The temporal separation supports the long-standing distinction between decision utility and experienced utility specifically, and multiple representations of utility in general (*27*). More broadly, our results support the Kantian and Rawlsian intuition that justice is rooted in a sense of fairness; yet contrary to Kant and Rawls, such a sense is not the product of applying a rational deontological principle but rather results from emotional processing, providing suggestive evidence for moral sentimentalism.

References and Notes

1. Aristotle, W. D. Ross, J. O. Urmson, *The Nicomachean Ethics* (Oxford Univ. Press, Oxford, 1980).
2. L. Kohlberg, *The Philosophy of Moral Development: Moral Stages and the Idea of Justice* (Harper and Row, San Francisco, CA, ed. 1, 1981).
3. A. K. Sen, *Collective Choice and Social Welfare* (Elsevier, New York, 1984).
4. J. C. Harsanyi, *Essays on Ethics, Social Behavior, and Scientific Explanation* (D. Reidel, Dordrecht, Holland, 1976).
5. J. Rawls, *A Theory of Justice* (Clarendon, Oxford, 1972).
6. A. Smith, *The Theory of Moral Sentiments* (Longman, Hurst, Rees, Orme, and Brown, London, ed. 11, 1812).
7. D. Hume, *A Treatise of Human Nature* (Clarendon, Oxford, ed. 2, 1978).
8. J. D. Greene, R. B. Sommerville, L. E. Nystrom, J. M. Darley, J. D. Cohen, *Science* **293**, 2105 (2001).
9. A. G. Sanfey, J. K. Rilling, J. A. Aronson, L. E. Nystrom, J. D. Cohen, *Science* **300**, 1755 (2003).
10. D. Engelmann, M. Strobel, *Am. Econ. Rev.* **94**, 857 (2004).
11. M. R. Delgado, R. H. Frank, E. A. Phelps, *Nat. Neurosci.* **8**, 1611 (2005).
12. E. Fehr, M. Naef, K. M. Schmidt, *Am. Econ. Rev.* **96**, 1912 (2006).
13. W. Schultz, *Nat. Rev. Neurosci.* **1**, 199 (2000).
14. J. P. O'Doherty, *Curr. Opin. Neurobiol.* **14**, 769 (2004).
15. D. J.-F. de Quervain *et al.*, *Science* **305**, 1254 (2004).
16. W. T. Harbaugh, U. Mayr, D. R. Burghart, *Science* **316**, 1622 (2007).
17. F. Krueger *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 20084 (2007).
18. J. Moll *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 15623 (2006).
19. T. Singer *et al.*, *Nature* **439**, 466 (2006).
20. T. Singer *et al.*, *Science* **303**, 1157 (2004).
21. A. B. Atkinson, *J. Econ. Theory* **2**, 244 (1970).
22. Y. Amiel, F. A. Cowell, *Thinking About Inequality: Personal Judgment and Income Distributions* (Cambridge Univ. Press, Cambridge, 1999).
23. See supporting data on *Science* Online.
24. C. M. Kuhnen, B. Knutson, *Neuron* **47**, 763 (2005).
25. K. Preuschoff, P. Bossaerts, S. R. Quartz, *Neuron* **51**, 381 (2006).
26. We assume the utility function for participant i to be $u_i(x) = \sum_{e \in x} x - \alpha_i \cdot \text{gini}(x)$, where x is the vector of allocations. Efficiency is the total number of meals donated, and inequity is measured by the "gini" coefficient. The parameter α captures the weighting placed upon inequity. We denote the chosen total meals as M_C, the unchosen meals M_U, and the marginal number of meals $\Delta M = M_C - M_U$. Terms for gini (G) and utility (U) are defined similarly (*22*).
27. D. Kahneman, P. P. Wakker, R. Sarin, *Q. J. Econ.* **112**, 375 (1997).
28. Although we controlled for gender and age effects in the regression model, we cannot exclude the hypothesis that this relation is driven in part by gender effects (*22*). Prominent gender difference in fairness behavior has been reported, among others, in (*12*).
29. P. R. Montague, T. Lohrenz, *Neuron* **56**, 14 (2007).
30. R. Loeber, J. D. Burke, B. B. Lahey, A. Winters, M. Zera, *J. Am. Acad. Child Adolesc. Psychiatry* **39**, 1468 (2000).

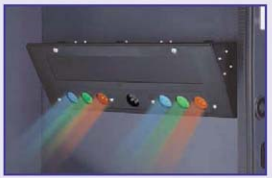
31. This project was supported by the David and Lucile Packard Foundation, the Gordon and Betty Moore Foundation, the John Templeton Foundation (S.R.Q.), and the Beckman Institute (M.L.H.). We thank the Canaan Children's Home for their cooperation; M. Stewart for research assistance; and T. Anastasio, K. Binmore, P. Bossaerts, C. Camerer, P. Glimcher, S. Mysore, R. Montague, and S. Williams for valuable suggestions and discussion of ideas.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1153651/DC1
Materials and Methods
SOM Text
Figs. S1 to S7
Tables S1 to S14
References
Movies S1 and S2

3 December 2007; accepted 21 April 2008
Published online 8 May 2008;
10.1126/science.1153651
Include this information when citing this paper.

New Products



Fluorescent Dye Imaging

The RGB Lighting Module for the G:BOX chemiluminescence imaging and documentation systems is designed to provide the flexibility to quickly image gels or blots stained with commercially available fluorescent and visible dyes. The module has 84 high-intensity LEDs providing red, blue, green, and white light. Additional slots in the module also allow the fitting of ultraviolet tubes emitting light of 365 nm or 254 nm. The module, which can be fitted to the complete range of G:BOX chemiluminescence systems, is fully computer-controlled, as is the filter wheel.

Syngene
For information 800-686-4407
www.syngene.com

Multimode Microplate Reader

The Infinite M1000 offers flexibility, sensitivity, and speed for a wide range of detection modes, and has been optimized for time-resolved fluorescence energy transfer assays as well as glow luminescence, fast luminescence, and dual-color luminescence assays. The Infinite M1000 provides the sensitivity and speed typically associated with filter-based systems. The platform can be upgraded with new detection modes. The system provides high-quality performance for a broad spectrum of applications, including the latest biomolecular assays for primary and secondary screening, receptor-ligand binding studies, cell-based assays, and ultraviolet fluorometry-based applications. The Infinite M1000 has an optional state-of-the-art injector module for up to two injectors, and can be combined with a stacker module or fully integrated with the Freedom Evo series of automated liquid-handling workstations.

Tecan
For information +41 44 922 81 11
www.tecan.com

Silica-Like Beads

Two new kits, Dynabeads Silane viral NA and Dynabeads Silane genomic DNA kits support isolation of low titers of viral RNA and DNA and pure genomic DNA of high integrity, respectively. Dynabeads provide highly predictable and reproducible isolation of total nucleic acids from biological samples. These validated magnetic beads can be used for both high-sensitivity in vitro diagnostic assays and high-capacity applications. They are uniform in size range distribution, surface area, and binding capacity to provide consistent assay reproducibility and results.

Invitrogen
For information 800-955-6288
www.invitrogen.com

Fast Raman Imaging

SWIFT and DuoScan are two new fast Raman scanning technologies. SWIFT (scanning with incredibly fast times) enables measurement times so fast that a 50,000 spectrum image can be obtained in six minutes. It can be used with a choice of detectors and makes use of either standard motorized mapping stages or piezo stages for use when very small step sizes are required. DuoScan introduces a method in which scanning hardware enables the image pixel size to be chosen to match features of interest, enabling optimized mapping speeds. Unlike many Raman imaging methods, the DuoScan retains the full confocal advantage of the Raman technique without restrictions on laser wavelength. Combining SWIFT and DuoScan can result in unprecedentedly fast mapping of large samples.

Horiba Jobin Yvon
For information 800-446-7422
www.horiba.com

Sequence Analysis Software

A new suite of "Next Generation" software for sequencing systems features single nucleotide polymorphism mutation analysis and expression, target sequence, and de novo assembly modules for data generated by the Illumina Genome Analyzer (Solexa), 454 Genome Sequence FLX, and Applied Biosystems SOLiD sequencing systems. The programs are designed to deliver rapid assembly and accurate analysis of the vast amounts of information generated by these instruments.

SoftGenetics
For information 814-237-9340
www.softgenetics.com

Endothelial Cell Invasion Kit

The Cultrex Endothelial Cell Invasion Kit for measuring early angiogenesis in vitro is designed to accelerate the screening process for compounds that influence vascular endothelial cell invasion and migration across the basement membrane. The kit offers a flexible, standardized, high-throughput format for quantitating the degree to which endothelial cells penetrate a barrier of basement membrane extract, in vitro, in response to chemoattractants or inhibiting. The assay features a simplified Boyden chamber-like design with a basement membrane-coated 8-micron polyethylene terephthalate membrane.

Trevigen
For information 800-873-8443
www.trevigen.com

DNA from Blood

The versatile UltraClean Blood DNA Isolation Kit can isolate DNA from whole blood in samples ranging from 300 microliters to 1 liter. The kit is based on a unique way of separating lymphocytes and lysing them to recover pure genomic DNA. The DNA is ready for applications such as polymerase chain reaction, restriction digest, and microarrays. This kit can also be used on other body fluids such as semen, saliva, buccal cells, and bone marrow.

Mo Bio Laboratories
For information 800-606-6246
www.mobio.com

Electronically submit your new product description or product literature information! Go to **www.sciencemag.org/products/newproducts.dtl** for more information.

Newly offered instrumentation, apparatus, and laboratory materials of interest to researchers in all disciplines in academic, industrial, and governmental organizations are featured in this space. Emphasis is given to purpose, chief characteristics, and availabilty of products and materials. Endorsement by *Science* or AAAS of any products or materials mentioned is not implied. Additional information may be obtained from the manufacturer or supplier.